N6739049

A PREVIEW CONTROL MODEL WITH ONE OR TWO FAST-TIME SCALE LOOPS

by

Richard Allen Miller

B.M.E., Union College
(1966)

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE
DEGREE OF MASTER OF
SCIENCE
at the
MASSACHUSETTS INSTITUTE OF
TECHNOLOGY
September, 1967

Signature of Author..... R. Allen Miller .....
Department of Mechanical Engineering, August 21, 1967

Certified by..... Thomas B. Sheridan .....
Thesis Supervisor

Accepted by.....
Chairman, Departmental Committee
on Graduate Students

REPRODUCED BY:
U.S. Department of Commerce
National Technical Information Service
Springfield, Virginia 22161
NTIS

A PREVIEW CONTROL MODEL

WITH ONE OR TWO FAST-TIME SCALE LOOPS

by

Richard Allen Miller

Submitted to the Department of Mechanical Engineering
on August 21, 1967, in Partial Fulfillment of
the Requirement for the Degree of
Master of Science

This thesis was concerned with development of a model of the human operator in control situations where the human can view the future input. The modeling technique used either one or two dynamic models of the controlled plant, operating in time scales faster than real time, to produce an estimate of the future response of the plant. This simulated the human's ability to estimate future responses.

The preview problem was treated as an optimization problem and solved analytically via the calculus of variations. It was shown that the use of fast-time predictor models provided a stable method of numerically solving the resultant two point boundary value problem. It was also demonstrated that an increase in the number of fast predictor models had the effect of additional iterations in a numerical approach to the solution of the two point boundary value problem. Additional fast models resulted in better future response estimation with the final system output approaching the optimum.

The various parameters and possible configurations of the system were discussed. General statements were made concerning the best configuration and the way in which preview should be used.

Extensive computer simulations were performed and the results for step and continuous inputs reported. Two human subjects were tested and their average responses matched by the preview control model discussed in this thesis. For the plants used the match was of good quality.

Thesis Supervisor: Thomas B. Sheridan
Title: Associate Professor of Mechanical Engineering

ACKNOWLEDGEMENTS

The author expresses his appreciation for the opportunity of working under the guidance of his thesis advisor, Prof. Thomas B. Sheridan. His earlier work with preview control models was the stimulus for this project.

Several other members of the Man-Machine Systems Lab. have been of assistance to the author. These include especially Prof. W. R. Ferrell and W. H. Vickers. Vickers' many comments have been extremely useful and his work with preview control showed that an optimization problem could be applied to the preview situation. Vickers' work was the foundation for the analysis done in this thesis.

This thesis was supported in part by the N.A.S.A. through the Division of Sponsored Research of Massachusetts Institute of Technology, DSR number 70283. The author was supported by a N.S.F. fellowship.

TABLE OF CONTENTS

		PAGE
I.	INTRODUCTION	1
II.	HISTORY OF MODEL DEVELOPMENT	2
III.	OBJECTIVES OF THE STUDY	3
IV.	ANALYSIS OF THE PREVIEW SITUATION	9
V.	SIMULATION RESULTS FROM TWO-TIME SCALE MODEL	19
VI.	TWO FAST-TIME MODELS	23
VII.	SIMULATION RESULTS FROM TWO FAST MODEL APPROACH	28
VIII.	OPTIMUM PREVIEW SPAN AND SAMPLE PERIOD	34
IX.	NON OPTIMUM WEIGHTING FUNCTIONS	39
X.	DESIGN OF THE FIRST FAST MODEL	42
XI.	PERFORMANCE WITH CONTINUOUS INPUTS	46
XII.	DISCONTINUOUS IMPORTANCE WEIGHTING FUNCTIONS	48
XIII.	HUMAN SUBJECT DATA	52
XIV.	CONCLUSIONS AND RECOMMENDATIONS	56
	REFERENCES	60
APPENDIX A	BRIEF DESCRIPTION OF THE KELLEY PREDICTOR INSTRUMENT	61
APPENDIX B	COMPUTER PROGRAM DESCRIPTIONS AND FLOW CHARTS	63
APPENDIX C	DATA	69
APPENDIX D	EXAMPLES OF OPTIMUM WEIGHTING FUNCTION CALCULATION	94

LIST OF FIGURES

2.1 Preview control characteristics

2.2 Held value preview model

2.3 One fast model preview control system

4.1 Typical preview situation

4.2 Plant description

4.3 Typical error weighting function

5.1 Demonstration of shifting preview interval

5.2 Optimum error weighting function for single integrator

5.3 Response, $\frac{1}{s}$, one fast model

5.4 Preview vs. performance index, $\frac{1}{s}$

6.1 Two fast-time model preview control system

7.1 Response, $\frac{1}{s}$, two fast models

7.2 Typical response, two fast models

7.3 Typical first fast model output

7.4 Typical first fast model output

7.5 Typical first fast model output

7.6 Typical second fast model output

8.1 Presumed response estimation error vs. preview time for human

8.2 Response estimation error vs. preview time for computer model

9.1 Decreasing error weighting function (arbitrary)

9.2 Linearly decreasing error weighting function

10.1 Simplest configuration for first fast model

10.2 Compensated fast model configuration

12.1 Human subject data with discontinuous inputs and non-uniform importance

13.1 Apparatus for human subject data

A-1 Kelley's predictor display for human operators

B-1 Flowchart of simulation program for preview control model using two fast models

C-1 Response, $\frac{1}{s(s+1)}$

C-2 Preview time vs. performance index $\frac{1}{s(s+1)}$

C-3 Response, $\frac{1}{(s+1)}$

C-4 Response, $\frac{1}{s(2s+1)}$

C-5 Preview time vs. performance index $\frac{1}{s(2s+1)}$

C-6 Response, $\frac{1}{s(2s+1)}$, T = 1.5

C-7 Response, $\frac{1}{s(2s+1)}$, T= 4.5

C-8 Response, $\frac{5}{s(3s+1)}$

C-9 Preview time vs. performance index, $\frac{5}{s(3s+1)}$

C-10 Response, $\frac{5}{s(3s+1)}$

C-11 Preview vs. performance index, $\frac{5}{s(3s+1)}$

C-12 Preview vs. performance index, comparison one and two fast models

C-13 Sample period vs. performance index

C-14 Response, $\frac{1}{s}$

C-15 Preview time vs. performance index, $\frac{1}{s}$

C-16 Response, $\frac{5}{s(3s+1)}$

C-17 Response. $\frac{5}{s(3s+1)}$

C-18 Preview time vs. performance index for non optimum error weighting

C-19 K_m vs. performance index

C-20 Closed loop poles, first fast model

C-21 Response, $\frac{5}{s(3s+1)}$, continuous input

C-22 Response, $\frac{5}{s(3s+1)}$, continuous input

C-23 Response, $\frac{5}{s(3s+1)}$, continuous input

C-24 Response, feedback controller, no preview, continuous input

C-25 Response, $\frac{5}{s(3s+1)}$, non-uniform importance

C-26 Comparison response with uniform importance

C-27 Response, $\frac{1}{s}$, non-uniform importance

C-28 Performance index vs. preview time, E varied

C-29 Human subject data, subject 1, $\frac{1}{s}$

C-30 Human subject data, subject 2, $\frac{1}{s}$

C-31 Human subject data, subject 1, $\frac{1}{s(s+1)}$

C-32 Human subject data, subject 2, $\frac{1}{s(s+1)}$

I. INTRODUCTION

A control system can be designed to operate in a wide variety of situations, and when a human is placed in the loop, the possibilities are even more numerous. The simplest type of control system consists of a loop which reacts to an input at an instant in time to produce an output at that instant. Humans performing instantaneous tracking functions fall in this class. When the future inputs can be previewed, with or without limitations, the system immediately becomes more complex and difficult to understand. If the rate at which the input is incurred is partially or wholly under the human operator's control, the task is said to be self-paced. In this case the independent input is actually a series of positions to be matched in the given available time. The alternative to self-pacing is forced-pacing in which input is received at a rate not under the control of the human.

In many tasks performed by humans, motions in two co-ordinate directions are interrelated in such a way as to make the task self-paced. We know from everyday experience that the human makes good use of his ability to see ahead or preview information control tasks.
For example, when driving an automobile one views the road and initiates control action to meet certain criteria, such as minimizing error, conserving fuel, minimizing time, or whatever. This task could be considered self-paced with constrained preview. The preview is constrained unless the input is known all the way to the desired destination.

Generally, when a human has preview he can perform a control task better than he can without it. Since the ability to preview is such an important feature of most real-life control tasks, it should be an equally important feature of a model of the human operator appropriate to such tasks. The preview control model studied in this thesis incorporates such a capability, and attempts to study how and why preview is so important. We shall restrict attention to the forced-paced situation, realizing the importance of the self-paced problem, but leaving it for future development.

II. HISTORY OF MODEL DEVELOPMENT

The basic reason for developing this preview model was to study how the human uses previewed information to his advantage in control tasks. The fundamental idea for such a model was found in a paper by Ziebolz and Paynter.[1] Their paper presented an idea for an automatic control system in which a computer model of the controlled plant or process, with time constants scaled to faster time than the actual system, is used to generate information about the future state of the controlled variables.

Kelley[2] adapted this principle in developing his predictor instrument for manual control. This instrument provides the operator with a prediction of the future response of the controlled plant based on the present state and forcing of the plant. (see appendix A)

Sheridan used the Ziebolz-Paynter-Kelley two-time scale predictor as a basis for a preview model of the human.[3] The essential idea is to use the high speed predictor to characterize the thought experiment performed by a human in anticipating his future response in a typical preview situation.

Consider the often used example of driving an automobile. The thought experiment might proceed as follows:

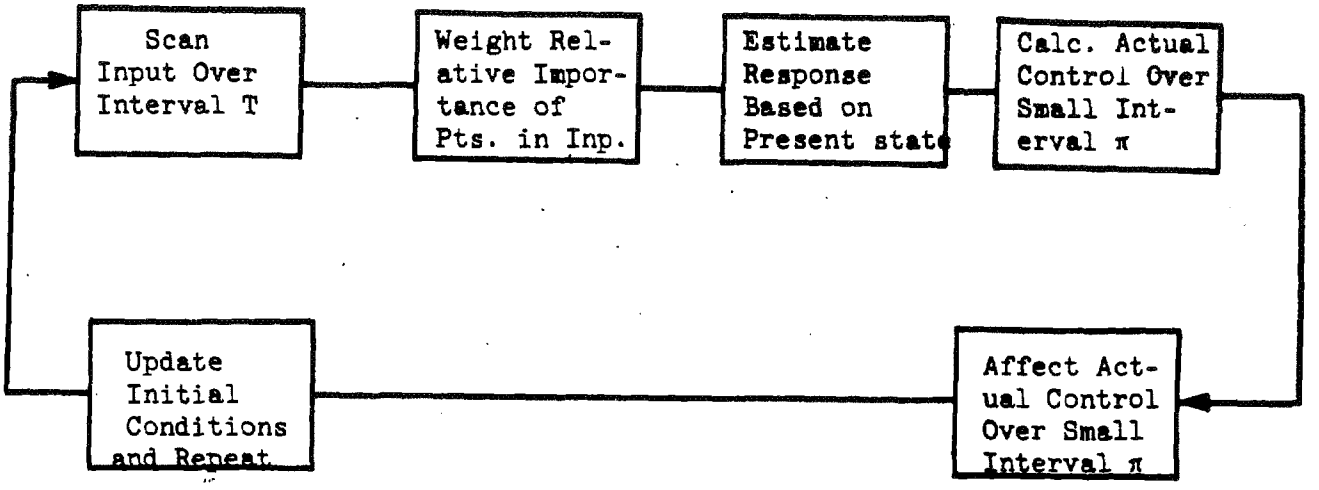
First, the operator scans the immediate roadway, noticing areas of possible difficulty such as sharp curves, bridges, parked cars, etc. At points of such possible difficulty, he puts an increased importance on the input effectively tightening the tolerances on his performance at these points. He now has pictured for himself a weighted path of travel.

Second, from his knowledge of the characteristics of the vehicle, he estimates the necessary control action to follow the input path. Presumably, this control is based on a weighted average of his anticipated response.

Third, after proceeding a short way he takes another look and decides on his next control action based on this updated information.

The basic ideas are characterized in figure 2.1.

If we use the predictor to model the anticipation process of the human, a diagram of such a model can be drawn. One type is shown in


Scan Input Over Interval T
Weight Relative Importance of Pts. in Inp.
Estimate Response Based on Present state
Calc. Actual Control Over Small Interval π
Update Initial Conditions and Repeat
Affect Actual Control Over Small Interval π

Preview control characteristics Figure 2.1

figure 2.2.

Assuming a constrained preview of the input for T units ahead, and assuming the sensors are capable of recognizing input positions with perfect accuracy, the system uses the fast-time dynamic model of its own characteristics to predict the future response of the system during the next T seconds.

The operation proceeds as follows:

1. The fast-time dynamic model of the process is initialized with the state variables of the process at time t_o.
2. The model is released with no change in control from the initial.
3. The output from the predictor is compared with the scanned input and weighted by the importance weighting function I(t), which is dictated by the environment.
4. The error is then weighted by the systems own error weighting function which indicates the degree of attention paid to various portions of the future predicted error.
5. The weighted error is integrated over the preview span T to determine the value of control for the next cycle.
6. When the next sampling point is reached, the fast dynamic model is reinitialized, and the cycle starts again.

The decay in front of the predictor is necessary to make the system operate like a good controller. If the system is a good one, the amount of control needed should decrease with time.

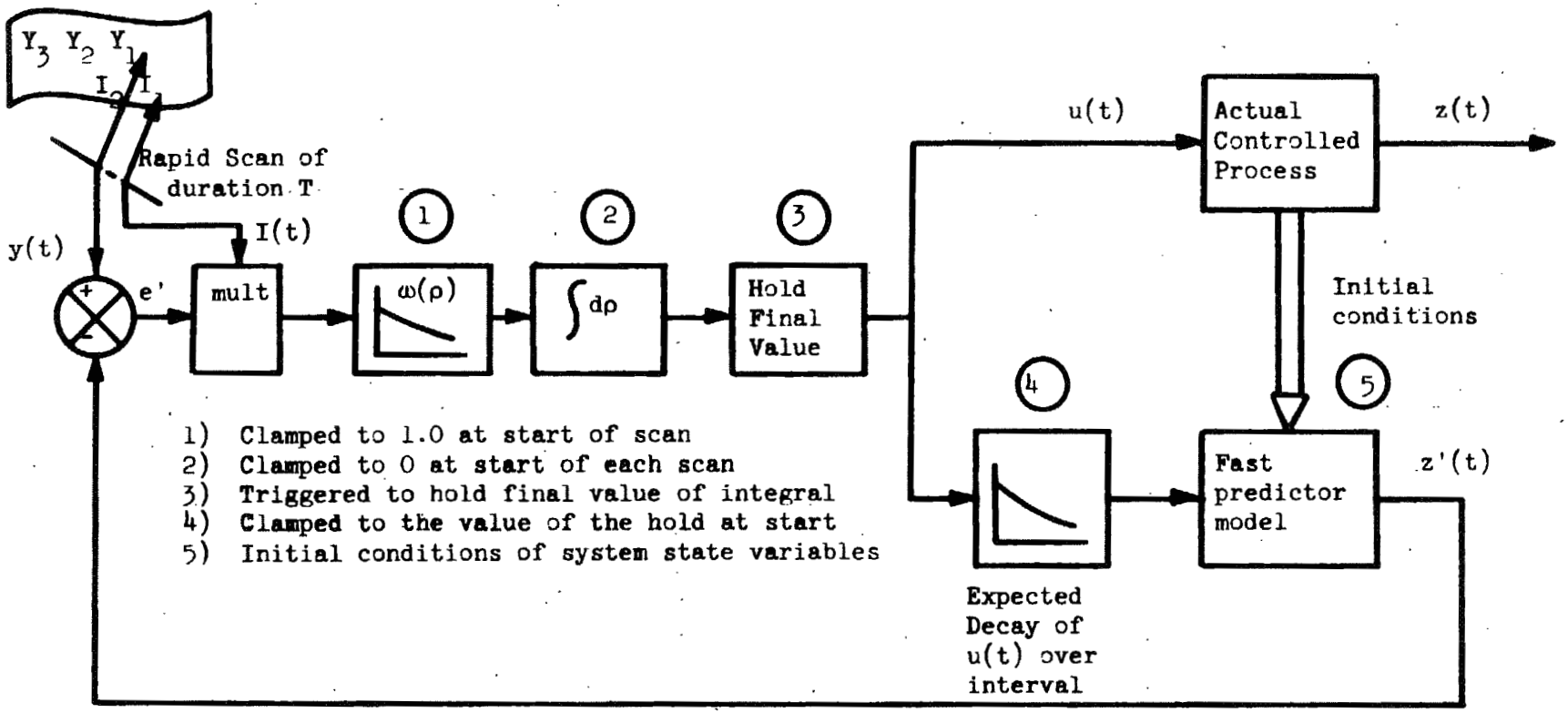


Figure 2.2 Held value preview model

This model was empiricallly studied via analog and digital simulation in a thesis by W. M. Johnson.[4] His results indicate that such a model does in fact provide a reasonable representation for the human operator in a preview situation.

A second slightly different model was also proposed in reference 3. A block diagram of this system is shown in figure 2.3.

In this case, the predictor model is placed in a feedback loop with the model forced in fast-time by the previewed input. The output of this model is again weighted by the importance weighting function and system error weighting function. The weighted future error is integrated over the preview span to determine the control for the next cycle.

A more mathematical discription is as follows:
The plant is characterized by its state equations.

$$\begin{aligned} \dot{\underline{x}} &= \underline{A}\underline{x} + \underline{B}\underline{u} \\ \underline{z} &= \underline{C}^T\underline{x} \end{aligned} \qquad 2.1$$

The fast dynamic model is characterized by the equations

$$\begin{aligned} \dot{\underline{x}}' &= \underline{A}'\underline{x}' + \underline{B}'\underline{u}' \qquad \text{time scale } \alpha \\ \underline{z}' &= (\underline{C}^T)'\underline{x}' \qquad \text{scale factor } k = \alpha/t \end{aligned} \qquad 2.2$$

with transition matrix $\underline{\Phi}(t,t_o)$ for prediction of $\underline{x}$ from any given $\underline{x}(t_o)$. The difference between equations 2.1 and 2.2 is the time scale, with 2.2 operating k times as fast as the controlled plant.

At time t_o, the fast-time model is initialized with the state variables of the plant. That is,

$$\underline{x}'(t_o) = \underline{x}(t_o) \qquad 2.3$$

Hence,

$$\underline{x}'(\alpha) = \underline{\Phi}(\alpha,t_o)\,\underline{x}(t_o) + \int_{t_o}^{\alpha} \underline{\Phi}(\alpha,\rho)\,\underline{B}\,u(\rho)\,d\rho \qquad 2.4$$

$$t_o \leq \alpha \leq t_o + T$$

where $$u(\rho) = K_m\, e(\rho) \qquad 2.5$$

The output is weighted by the importance and error weighting functions, and integrated to give the control.

$$u(t_o + \pi) = \int_{t_o}^{t_o + T} K_g\, \omega(\rho)\, e(\rho)\, d\rho \qquad 2.6$$

where π is the sample period and $e(\rho)$ is given by

$$e(\rho) = y(\rho) - \underline{C}^T \underline{x}'(\rho) \qquad 2.7$$

$y(\rho)$ being the input to the system. When real time reaches $t_o + \pi$, the fast-time model is initialized again with the state variables $\underline{x}(t_o + \pi)$ and the process repeats.

Johnson[4] performed some simulation studies of this model, and Sheridan[3] did some further digital studies of the system. Their results were encouraging enough to warrant further study, especially in the case of the second model, shown in figure 2.3. For these reasons, and for others which will become apparent later, attention will be restricted exclusively to this model.

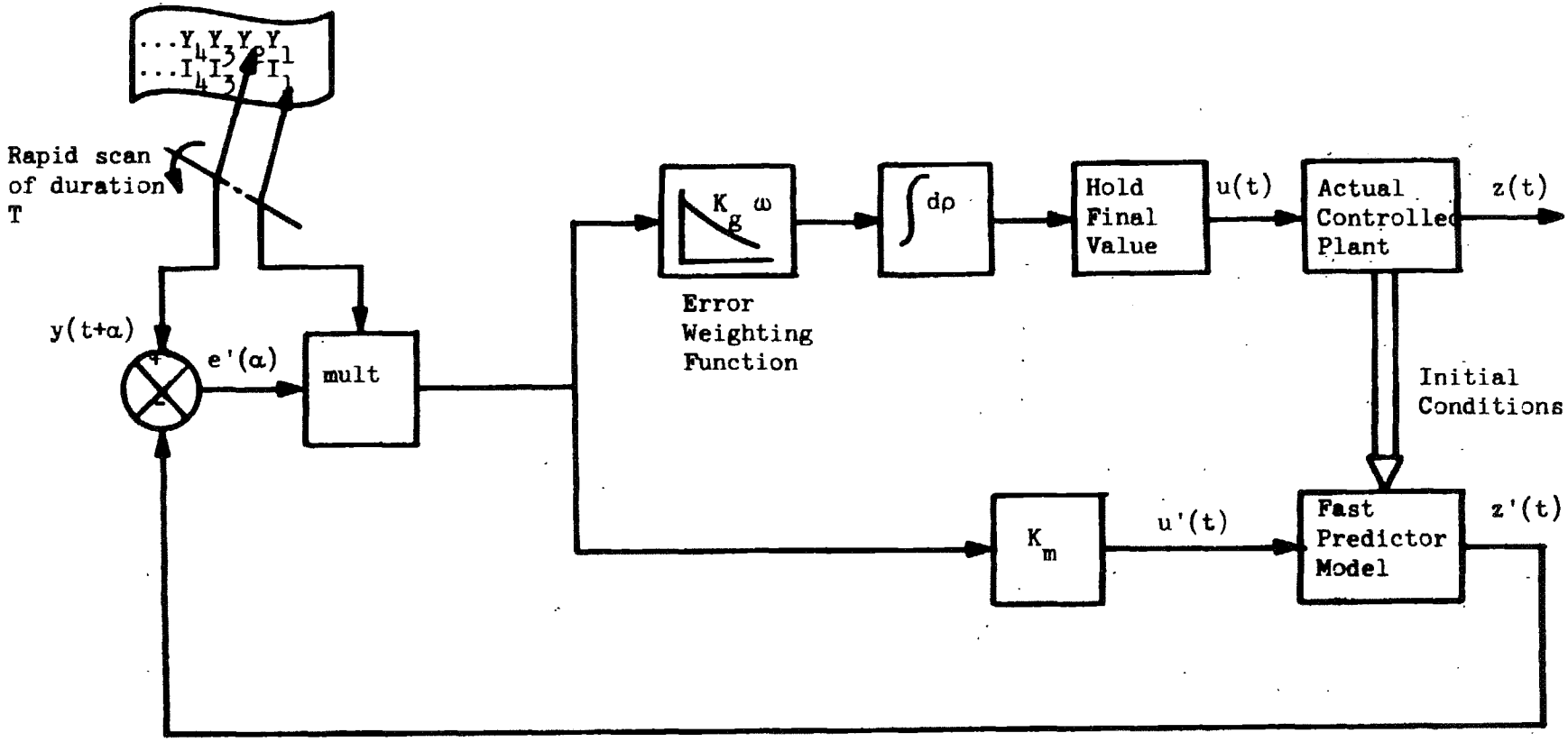


Figure 2.3 One fast model preview control system

III. OBJECTIVES OF THE STUDY

As we said earlier, the basic reason for developing a preview model is to provide a method of studying the effects of preview on performance in control tasks. It is hoped that this model can provide some clues as to how the human uses preview and why it is so important.

Related problems are:

1. Optimum preview span for a given plant.
2. Degree of correspondence with the real plant required of the model. (fast-time model)
3. Form of the error weighting function and values for other parameters of the model.

At the outset, it was decided to use digital simulation due mainly to the problems associated with previewing future inputs in an analog simulation.

IV. ANALYSIS OF THE PREVIEW SITUATION

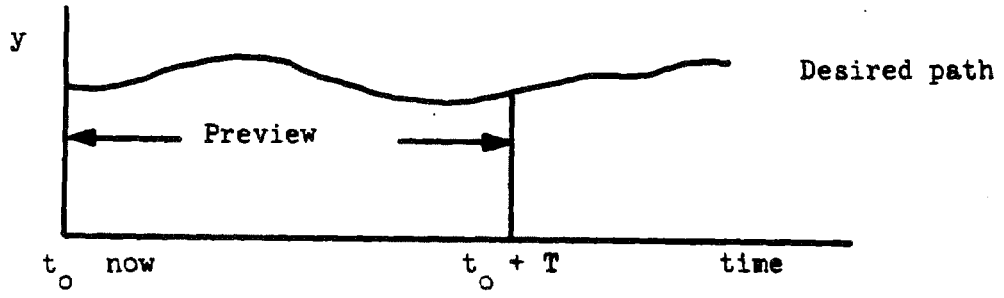
When one examines the proposed model (figure 2.3) one finds a minimum of four independent parameters for any given plant. They are:

1. Preview time
2. Sample period, or duration of the hold
3. Error weighting function
4. Gain K_m

The importance weighting function $I(t_o + \rho)$ is assumed known for a given input. It is obvious that straightforward trial and error experimentation would be near futile, and might never yield any truly significant information.

In an attempt to find some foundation upon which to base experiments, let us restrict the preview problem a little more, and set it up in terms of an optimization problem.*

As before, consider a control situation where the operator can see the desired trajectory for a finite distance into the future, as shown in figure 4.1.



Typical Preview Situation Figure 4.1

Now, suppose the operator wishes to traverse the path while minimizing some performance index. One method of approaching this problem would be to pick a control which optimizes over the span $(t_o, t_o + T)$, then travel a short way down the path and repeat the operation.

---

* W. H. Vickers, a co-worker in the Man-Machine Systems Lab. first noticed the analytic similarity between the two-time scale model and an optimum preview control situation.

This means the operator bases his control action for $(t_o, t_o + \pi)$ on the previewed input from $(t_o, t_o + T)$. When he reaches $t_o + \pi$ he again scans the input up ahead, this time to $t_o + \pi + T$ and tries to optimize his performance over this interval.

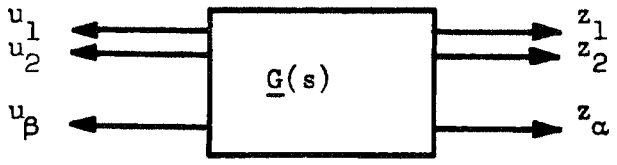
Except for the updating characteristic, the problem is the same as optimizing a system's performance subject to some index, with the desired state performance of the system known over a finite interval. This procedure is as follows:

We assume the plant is described by

$$\dot{\underline{x}} = \underline{A}\underline{x} + \underline{B}\underline{u} \qquad 4.1a$$

$$\underline{z} = \underline{C}^T\underline{x} \qquad 4.1b$$

with $\underline{x}(t_o)$ known. In block diagram form this corresponds to figure 4.2, with variables expressed in Laplace transform notation.



Plant Description Figure 4.2

$$\underline{G}(s) = \underline{C}^T(\underline{I}s - \underline{A})^{-1}\underline{B} \qquad 4.2$$

$$\begin{array}{ll} \underline{A} & n \times n \\ \underline{B} & n \times \alpha \\ \underline{u} & \alpha \times 1 \\ \underline{C} & n \times \beta \\ \underline{z} & \beta \times 1 \end{array} \qquad 4.3$$

We will assume a quadratic performance index of the form

$$\underline{H} = \int_{t_o}^{t_o + T} (\underline{e}^T\underline{E}\underline{e} + \underline{u}^T\underline{M}\underline{u}) \, d\rho \qquad 4.4$$

which in the scalar case becomes

$$H = \int_{t_o}^{t_o + T} (Ee^2 + Mu^2) \, d\rho \qquad 4.4a$$

with

$$\underline{e} = \underline{y} - \underline{C}^T \underline{x} \qquad 4.5$$

where $\underline{y}$ is the known input we are trying to follow, and $\underline{E}$ and $\underline{M}$ are positive definite matrices.

The objective is to minimize equation 4.4 subject to the constraint, equation 4.1a. Since $\underline{y}$ is known, the conditions of the problem imply that we determine a function $\underline{u}$ which satisfies equation 4.1, and minimizes $\underline{H}$.

The problem can be solved with the calculus of variations provided we meet certain restrictions.[4]

1. The elements of equation 4.1 are continuous with respect to $\underline{x}$ and $\underline{u}$.
2. They are continuously differentiable with respect to $\underline{x}$.
3. The controlled elements are assumed controllable and observable, i.e. all state variables are measurable and is possible to excite every state of the controlled elements.

We will not consider the case where the magnitude of any variable is restricted. If we did, the problem could still be handled.

Before trying to solve the problem, it might be worthwhile to briefly describe the procedures involved in the use of the calculus of variations and the method of Lagrange multipliers. Everyone is familiar with the problem of determining the maximum or minimum of a function. For example, a necessary condition that a continuously differentiable function $f(x_1, x_2, \ldots, x_n)$ have a relative maximum or minimum is

$$df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \ldots + \frac{\partial f}{\partial x_n} dx_n = 0 \qquad 4.7$$

at that point.

If the n variables are independent, the n differentials can be independently set equal to zero. However, if they are not, and are related by N conditions the problem is more difficult.

A procedure which is often used to solve the problem is the method of Lagrange multipliers.[6] For example, consider a function

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz \qquad 4.8$$

subject to two constraints

$$\phi_1(x,y,z) = 0 \qquad 4.9a$$

$$\phi_2(x,y,z) = 0 \qquad 4.9b$$

Since three variables must satisfy two constraint conditions, 4.9a and 4.9b, only one variable can be independently set equal to zero.

We now introduce as many new variables (Lagrange multipliers) as we have constraints. Equations 4.9a and 4.9b imply

$$\frac{\partial \phi_1}{\partial x} dx + \frac{\partial \phi_1}{\partial y} dy + \frac{\partial \phi_1}{\partial z} dz = 0 \qquad 4.10a$$

$$\frac{\partial \phi_2}{\partial x} dx + \frac{\partial \phi_2}{\partial y} dy + \frac{\partial \phi_2}{\partial z} dz = 0 \qquad 4.10b$$

Multiply 4.10a by λ_1 and 4.10b by λ_2, and add them to 4.8.

$$\left(\frac{\partial f}{\partial x} + \lambda_1 \frac{\partial \phi_1}{\partial x} + \lambda_2 \frac{\partial \phi_2}{\partial x} \right) dx + \left(\frac{\partial f}{\partial y} + \lambda_1 \frac{\partial \phi_1}{\partial y} + \lambda_2 \frac{\partial \phi_2}{\partial y} \right) dy + \left(\frac{\partial f}{\partial z} + \lambda_1 \frac{\partial \phi_1}{\partial z} + \lambda_2 \frac{\partial \phi_2}{\partial z} \right) dz = 0 \qquad 4.11$$

Now, we determine λ_1 and λ_2 so that two sets of the bracketed terms vanish. This means each individually must go to zero since one variable is independent. This leaves five equations, the three bracketed terms equal to zero and equations 4.9a and 4.9b, with five unknowns $x,y,z,\lambda_1,\lambda_2$.

The calculus of variations is also concerned with the determination of maxima and minima, or extrema. Rather than extrema of functions, however, the object is to determine extrema of functionals, where a functional is essentially a function of a function rather than of a variable. For example, $f(x(t))$ is a functional if, for each function $x(t)$ there corresponds a value f.

The procedure for calculating the extrema of functionals is analagous

to calculating the differential of a function. For example,

$$F = F(x, y(x), \frac{dy(x)}{dx}) \tag{4.12}$$

which for a fixed value of x depends on y and $\frac{dy}{dx}$.

The variation of F is

$$\delta F = \frac{\partial F}{\partial x} \delta x + \frac{\partial F}{\partial y} \delta y + \frac{\partial F}{\partial y'} \delta y' \tag{4.13}$$

However, we are concerned with a fixed value of x, so its variation is zero, leaving only the second and third terms.

Notice the differential of a function is a first order approximation to the change in that function along a particular curve, while the variation of a functional is a first order approximation to the change of the functional from curve to curve.

With this brief description of the method we now tackle the problem at hand. We want to find the function $\underline{u}(t)$ which will minimize the performance index (equation 4.4) subject to the constraint equation (4.1). If we restrict attention to the time invariant case, we have a functional $\underline{H}(t, \underline{x}, \underline{u})$. Let us rewrite the input $\underline{y}$ as

$$\underline{y} = \underline{C}^T \underline{q} \tag{4.14}$$

where $\underline{q}$ is the desired state performance we are trying to obtain at each instant of time. Rewriting 4.4 we get

$$\underline{H} = \int_{t_o}^{t_o + T} ((\underline{q} - \underline{x})^T \underline{C}\, \underline{E}\, \underline{C}^T (\underline{q} - \underline{x}) + \underline{u}^T \underline{M} \underline{u}) \, d\rho \tag{4.15}$$

$$H = \int_{t_o}^{t_o + T} F(t, x, u) \, d\rho$$

Taking a variation of 4.15 we get the following:

$$\delta \underline{H} = \int_{t_o}^{t_o + T} \left(\frac{\partial F}{\partial \underline{x}} \delta \underline{x} + \frac{\partial F}{\partial \underline{u}} \delta \underline{u} \right) d\rho \qquad 4.16$$

or

$$\delta H = \int_{t_o}^{t_o + T} \left((-2\underline{CEC}^T \underline{q} + 2\underline{CEC}^T \underline{x}) \delta \underline{x} + 2\underline{Mu}\, \delta \underline{u} \right) d\rho \qquad 4.17$$

Since $\underline{u}$ and $\underline{x}$ must satisfy equation 4.1, their variations cannot be arbitrarily assigned inside $(t_o, t_o + T)$. Therefore, the coefficients of $\delta \underline{x}$ and $\delta \underline{u}$ in 4.17 need not vanish independently. However, equation 4.1 implies

$$\delta \underline{\dot{x}} = \underline{A}\, \delta \underline{x} + \underline{B}\, \delta \underline{u} - \delta \underline{\dot{x}} = \underline{0} \qquad 4.18$$

If we multiply equation 4.18 by a Lagrange multiplier, which in general will be a time-varying 1 x n row vector, and integrate over $(t_o, t_o + T)$, we get the following equation.

$$\int_{t_o}^{t_o + T} (\underline{\lambda}^T \underline{A} \delta \underline{x} - \underline{\lambda}^T \delta \underline{\dot{x}} + \underline{\lambda}^T \underline{B} \delta \underline{u})\, d\rho = 0 \qquad 4.19$$

Now adding equations 4.19 and 4.17

$$\int_{t_o}^{t_o + T} \left((-2\underline{CEC}^T \underline{q} + 2\underline{CEC}^T \underline{x} + \underline{\lambda}^T \underline{A}) \delta \underline{x} - \underline{\lambda}^T \delta \underline{\dot{x}} + (\underline{\lambda}^T \underline{B} + 2\underline{Mu}) \delta \underline{u} \right) d\rho = 0 \qquad 4.20$$

Equation 4.20 must be true for any value of $\underline{\lambda}$.

Since $\frac{d\, \delta x}{dt} = \delta \left(\frac{dx}{dt}\right)$ we can integrate the middle term by parts, leaving

$$-\underline{\lambda}^T \delta \underline{x}\Big]_{t_o}^{t_o+T} + \int_{t_o}^{t_o+T} \left((-2\underline{CEC}^T\underline{q} + 2\underline{CEC}^T\underline{x} + \underline{\lambda}^T\underline{A} + \dot{\underline{\lambda}}^T)\delta \underline{x} + (2\underline{Mu} + \underline{B}^T\underline{\lambda}^T)\delta \underline{u}\right) d\rho \qquad 4.21$$

Since equation 4.21 is true for all $\underline{\lambda}$, we can choose $\underline{\lambda}$ such the coefficient of $\delta \underline{u}$ is zero. Then since $\delta \underline{x}$ can be arbitrarily assigned in the interval $(t_o, t_o + T)$, its coefficient must also vanish. Hence,

$$\dot{\underline{\lambda}}^T + \underline{\lambda}^T\underline{A} + 2\underline{CEC}^T(\underline{x} - \underline{q}) = \underline{0} \qquad 4.22a$$

$$\underline{B}^T\underline{\lambda}^T + 2\underline{Mu} = \underline{0} \qquad 4.22b$$

$$-\underline{\lambda}^T \underline{x}\Big]_{t_o}^{t_o+T} = \underline{0} \qquad 4.22c$$

Since $\underline{x}$ is known at t_o, $x(t_o) = \underline{0}$, leaving equation 4.21c as

$$-\lambda^T(t_o + T) = \underline{0} \qquad 4.22d$$

Substituting equation 4.22b into equation 4.1a, results in the following equations

$$\dot{\underline{x}}(t) = \underline{A}\underline{x}(t) + \underline{B}(-\tfrac{1}{2}\underline{M}^{-1}\underline{B}^T\underline{\lambda}^T) \qquad 4.23a$$

$$\dot{\underline{\lambda}}^T(t) = -\underline{\lambda}^T\underline{A} + 2\underline{CEC}^T(\underline{q} - \underline{x}) \qquad 4.23b$$

with boundary conditions

$$\underline{\lambda}^T(t_o + T) = \underline{0}$$

$$\underline{x}(t_o) = \underline{x}_o$$

Thus, we have 2n first order differential equations in 2n unknowns subject to n boundary conditions at t_o and n boundary conditions at $t_o + T$. Their simultaneous solution yields the optimum control $\underline{u}$. Notice that

equation 4.23b is the adjoint equation of the controlled process with system error $\underline{C}^T(\underline{q} - \underline{x})$ as the driving function.

In the ordinary case, equations 4.23a and 4.23b cannot be solved in closed form due to their coupling. Neither equation can be initiated at the appropriate time, i.e. t_o for equation 4.23a and $(t_o + T)$ for 4.23b, because neither $\underline{\lambda}^T(t_o)$ nor $\underline{x}(t_o + T)$ is known.

The most direct way of approaching this problem is to select a set of boundary conditions on a trial and error basis. However, a number of difficulties arise in the application of this method. First of all, the adjoint equation is unstable and makes the incidence of numerical errors difficult to avoid.

To see the reason for this, recall that the transition matrix of the adjoint equation is the transpose of the inverse of the actual system transition matrix. That is,

$$\underline{\Phi}_{adj}(t,t_o) = (\underline{\Phi}^{-1}(t,t_o))^T = \underline{\Phi}^T(t_o,t)$$

Hence, if the system is stable, the adjoint will be unstable.

However, in our case the situation is not so severe. Since the controller has a preview of future input, the output of the fast-time model, if it is a good model of the system, will be a good approximation to the future system response. Therefore, $\underline{\lambda}^T(t_o)$ can be found.

Solving equation 4.23b

$$\underline{\lambda}^T(t) = \underline{\Phi}^T(t_o,t)(\underline{\lambda}^T(t_o) + \int_{t_o}^{t} \underline{\Phi}^T(\rho,t_o)2\underline{CEe}(\rho)\, d\rho) \quad 4.24$$

where $\underline{e}$ is the system error as previously defined.

$\underline{\lambda}^T(t_o + T) = \underline{0}$ implies

$$\underline{\lambda}^T(t_o + T) = \underline{\Phi}^T(t_o,t_o + T)(\underline{\lambda}^T(t_o) + \int_{t_o}^{t_o + T} \underline{\Phi}^T(\rho,t_o)2\underline{CEe}(\rho)\, d\rho) = \underline{0} \quad 4.25$$

which upon simplification gives

$$\underline{\lambda}^T(t_o) = -\int_{t_o}^{t_o + T} \underline{\Phi}^T(\rho, t_o) 2\underline{CEe}(\rho)\, d\rho \qquad 4.26a$$

With both boundary conditions known for the adjoint equation, we can write down the solution to 4.23b. Substituting equation 4.26a into 4.24 we get the following:

$$\underline{\lambda}^T(t) = -\underline{\Phi}^T(t_o, t) \int_{t_o}^{t_o + T} \underline{\Phi}^T(\rho, t_o) 2\underline{CEe}(\rho)\, d\rho + \underline{\Phi}^T(t_o, t) \int_{t_o}^{t} \underline{\Phi}^T(\rho, t_o) 2\underline{CEe}(\rho)\, d\rho \qquad 4.27$$

or equivalently

$$\underline{\lambda}^T(t) = -\underline{\Phi}^T(t_o, t) \int_{t}^{t_o + T} \underline{\Phi}^T(\rho, t_o) 2\underline{CEe}(\rho)\, d\rho \qquad 4.27a$$

It is at this point that the instability becomes a problem since the term multiplying the integral in 4.27a is growing with time. Again we are saved by a characteristic of the model. We have postulated that only a short distance is traveled in real time before another calculation is made. Hence, if π is the sample period and $\pi \ll T$ for $t_o \leq \gamma \leq t_o + \pi$

$$\underline{\lambda}^T(t_o + \gamma) = -\underline{\Phi}^T(t_o, t_o + \gamma) \int_{t_o + \gamma}^{t_o + T} \underline{\Phi}^T(\rho, t_o) 2\underline{CEe}(\rho)\, d\rho \qquad 4.28$$

If we keep π small, equation 4.28 does not blow up.

With the equation for $\underline{\lambda}^T$, (4.28) we can now solve equation 4.21b for the optimum control between samples.

$$\underline{u}(t_o + \gamma) = \underline{M}^{-1}\underline{B}^T\underline{\Phi}^T(t_o, t_o + \gamma) \int_{t_o + \gamma}^{t_o + T} \underline{\Phi}^T(\rho, t_o)\underline{CEe}(\rho)\, d\rho \qquad 4.29$$

The two most significant points of the above are:

1. If the future response is accurately known, from the fast-time model, the optimum control can be generated.
2. Assuming (1.), if the sample period is short enough the present method of calculation generates a good approximation to the optimum.

The key word here is accurate. As we shall show in the next section, generation of the true optimum control depends on the fast-time model mimicking the actual system output over the preview time T.

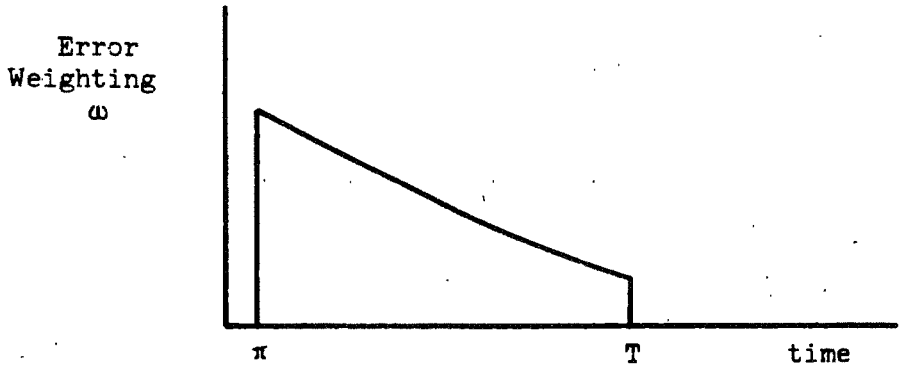
Equation 4.29 can be expressed in terms of the parameters of out basic model shown in figure 2.3. For this model, $\underline{M}$, $\underline{u}$, $\underline{E}$, and $\underline{e}$ are scalars and $\underline{C}$ and $\underline{B}$ are column vectors. The transition matrix is a matrix exponential. Hence, at the sample points the control is given by

$$u(t_o + \pi) = \frac{\underline{E}}{\underline{M}} \underline{B}^T \int_{t_o + \pi}^{t_o + T} e^{\underline{A}^T(\rho - t_o - \pi)} \underline{C}\, e(\rho)\, d\rho \qquad 4.30$$

Comparing with equation 2.6, we see that for the optimum system

$$K_g\, \omega(\rho) = \frac{\underline{E}}{\underline{M}} \underline{B}^T e^{\underline{A}^T(\rho - t_o - \pi)} \underline{C} \qquad 4.31$$

with the integration starting at $t_o + \pi$ instead of t_o, or equivalently, $\omega(\rho)$ is of the form shown in figure 4.3.



Typical error weighting function

Figure 4.3

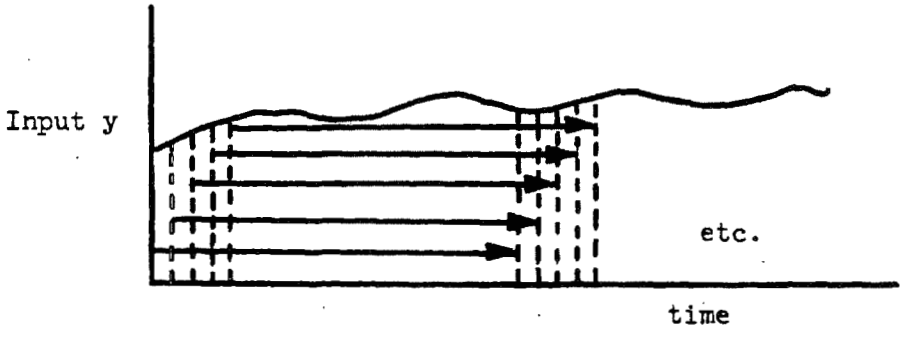
V. SIMULATION RESULTS FROM TWO-TIME SCALE MODEL

Computer programs were written for the model configuration shown in figure 2.3. (see appendix B for flow charts and methods) The general procedure was to vary the parameters, K_m, ω, and preview time T, to determine optimum performance in terms of an arbitrarily selected performance index

$$H = \int_0^{\infty} (Ee^2 + Mu^2)\, d\rho \qquad 5.1$$

Most of the data was taken with E =1, and M = 1, that is, weighting error and control effort equally.

A comment about the limits of equation 5.1 is in order. The optimization problem was solved over the interval $(t_o, t_o + T)$, but this interval is shifting in real time, as shown by figure 5.1. This means that a separate optimum control is computed over each interval T. Since each interval is optimum, the entire period in which an input is present $(0,\infty)$ will be under optimum control.



Demonstration of shifting preview interval — Figure 5.1

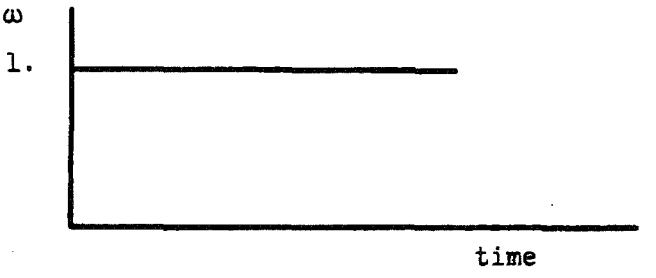
The first case studied was a plant with transfer function

$$G(s) = \frac{z(s)}{u(s)} = \frac{1}{s} \qquad 5.2$$

which in state form is

$$\frac{dx}{dt} = u. \qquad 5.3$$

Hence, A = 0, and $e^{At} = 1$. Refering to equation 4.31 the theoretical optimum error weighting function can be seen to be simply $\omega(t) = 1$, with K_g equal to unity. That is,



Opt. error weighting for $\frac{1}{s}$ Figure 5.2

Since the importance weighting function is an input to the system, any type could have been assumed. For convenience, all data was taken with uniform importance weighting, i.e. all errors equally important.

A typical step response for this system is shown in figure 5.3. Obviously, preview allowed the system to start its response prior to encountering the discontinuity. It was found that T = 1.5 produced the lowest performance index for this system. The value of H for T = 1.5 was not necessarily the lowest value of H possible for the given plant, but it was the lowest for this fixed control system with only preview time T as a variable.

Before exploring further characteristics of the model, it was decided to compare the predicted output, or output of the fast-time model, with the resulting system output. The results are shown as dotted curves in figure 5.3, where the dotted curves are weighted and integrated to determinte the control at the next sample point. These results were rather discouraging. At time 1.5, (real time) the system output was approximately 0.16 units. Starting with this initial condition, the predicted output over the preview span followed the dotted curve; far different from the actual system response over the same period. Farther on in time, say at 2.0, the response predicted more closely resembled that which actually happened.

This is to be expected. When one examines exactly what happens, he sees that the fast model is an error nulling servo, which tries to remove

error as fast as possible. When the initial conditions leave a negative error, as was the case at t = 1.5, the servo drives the output toward zero until the step is reached and a positive error results driving the input positive again.

It quickly becomes apparent that we are trying to model a more complex control action than that used in the Kelley predictor, or the held value model shown in Figure 2.2. The method used in those models assumed either the control stick returned to zero after an initial action, or that it held the same constant value for the entire period.

Since the model we are studying is more complex, it requires a more complex fast-time predictor model. When solving the optimum problem, we assumed a good knowledge of the dynamics, including the controller and required extrapolation of the future system output. Recall that the solution of equation 4.23b was possible only because a knowledge of the future system error was assumed. This means the fast model must anticipate the resulting system output, which in turn depends on the output of this same fast-time predictor. As we have seen above, the servo model does not give a good estimate of the actual resulting system response. Hence, the assumptions made earlier in solving the optimum problem are violated and the performance index has not been minimized.

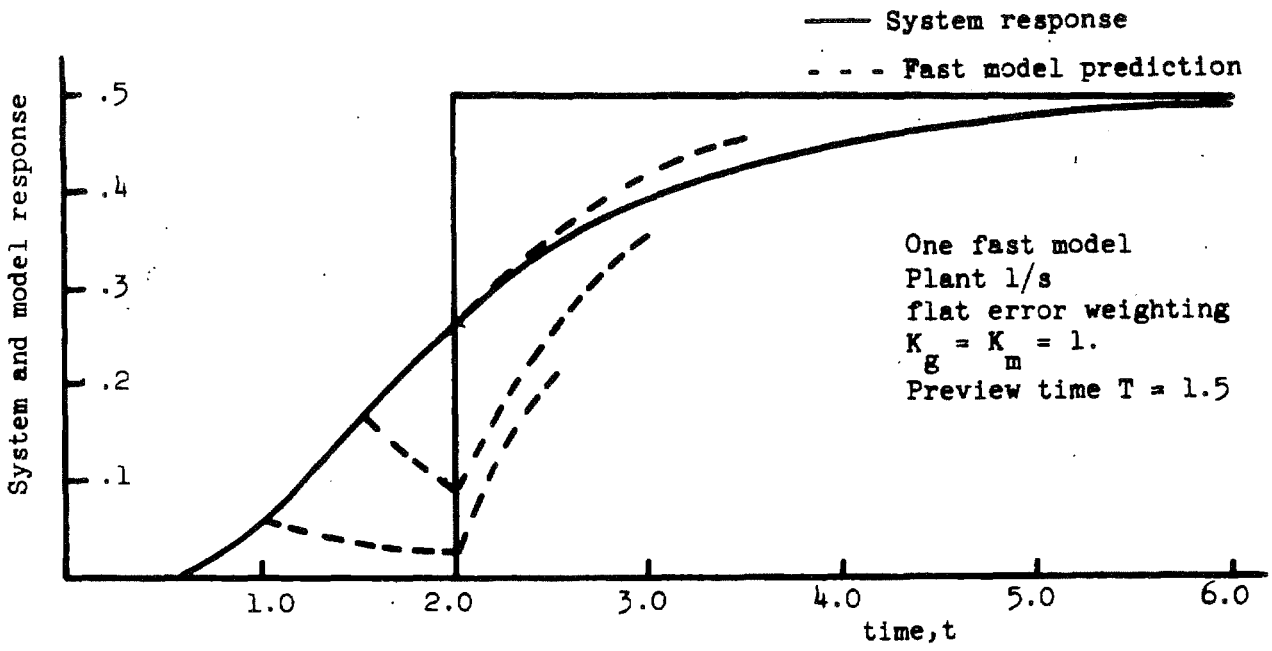


Figure 5.3

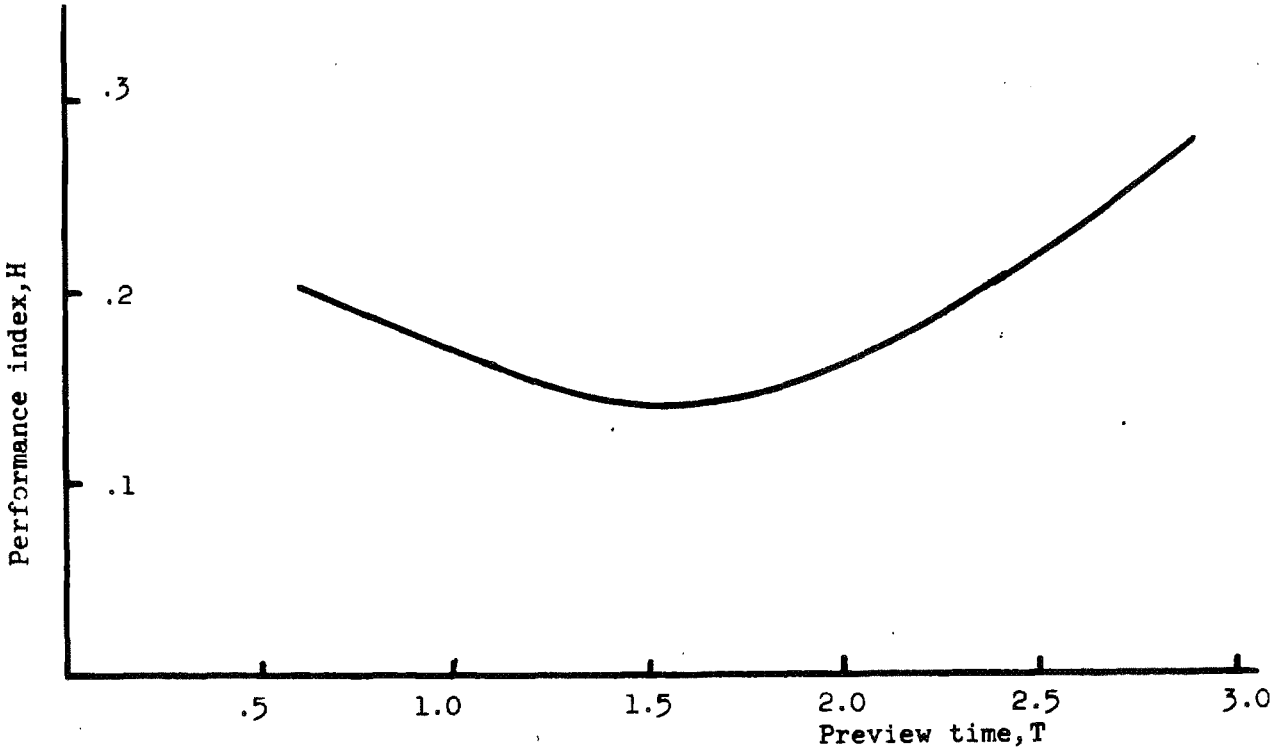


Figure 5.4

VI. TWO FAST-TIME MODELS

Recalling the results stated in section IV concerning solution of the general two point boundary value problem, a method of reducing the magnitude of the above difficulty becomes apparent. The standard boundary condition iteration method requires an additional guess at the magnitude of the boundary conditions and recalculation of the control if the results of the first guess are not satisfactory. When applied to our situation, this implies including at least one, and possibly more, additional fast-time scale loops to provide a better estimate of the future output. By incorporating more models, we can improve the accuracy of the estimate of the future response and approach the true minimum value of the performance index as close as desired.

On an intuitive basis, what does this mean? In the earlier theoretical formulation of the system, we assumed the fast-time model was an accurate replica of the actual controlled process behavior. However, the simplest feedback servo loop or even a simple loop with some compensating dynamics has no inherent anticipatory character. Since anticipation is the key feature of the system, it is reasonable to expect the fast model itself to have the same type of preview available as does the overall system. The addition of another fast model provides exactly that.

Consider the block diagram shown in figure 6.1. There are two model loops, the fastest of which we will call the first fast model, the slower we will call the second fast model. Thus, there are three time scale levels in the system. First, real time t, corresponding to the actual plant. Second, α, the time scale of the second fast model. The time scale of the first fast-time model, and the fastest time scale, is called β.

The method of operation is essentially the same as the two-time scale model described previously. Let us verbally walk through one cycle. At time t_o, both the first and second fast models are initialized with the state variables and control of the plant. Between time t_o and $t_o + \mu$, where μ is the sample period of the first fast model, the first fast-time model estimates the future output, and therefore future error, for the period $(t_o, t_o + T)$. From this, the control for the second model at $t_o + \mu$ is calculated. At $t_o + \mu$, the first model is reinitialized and it pro-

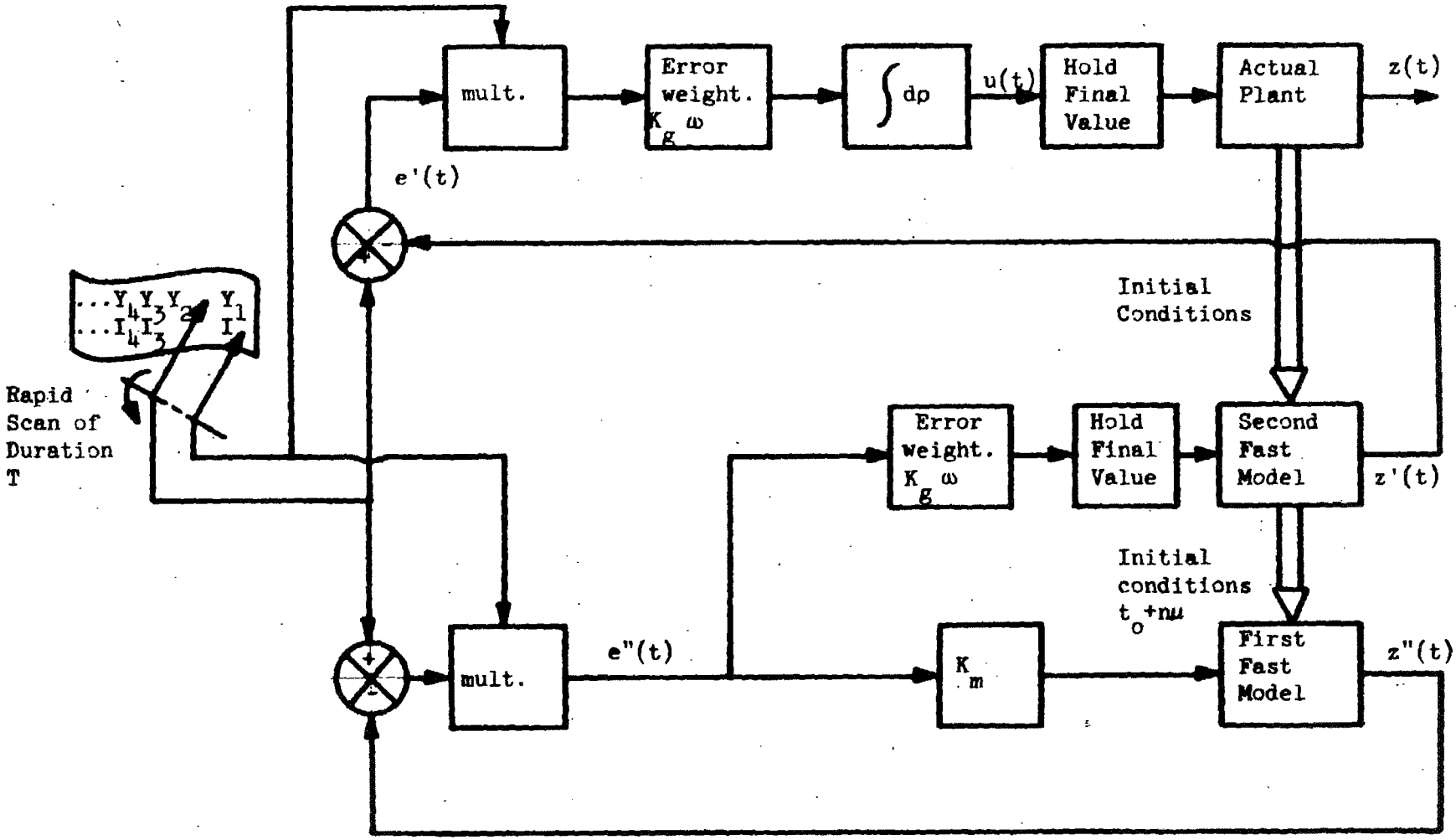


Figure 6.1 Two fast-time model preview control system

ceeds to estimate the error for the period $(t_o + \mu, t_o + T)$. Control for the second model at $t_o + 2\mu$ is then calculated. This continues until the second model reaches $t_o + T$. The trajectory of the second model, having been forced at $t_o + \mu$, $t_o + 2\mu$, etc., is a second and better estimate of the future response from $(t_o, t_o + T)$. From this, the actual control for the plant is calculated for time $t_o + \pi$. Then the whole cycle repeats with t_o replaced by $t_o + \pi$ and preview extending to $t_o + \pi + T$.

A more mathematical description is as follows:
Except for time scaling, the state equations for all three models are the same. Namely,

$$\dot{\underline{x}} = \underline{A}\underline{x} + \underline{B}\underline{u}$$

$$\underline{x} = e^{\underline{A}(t - t_o)}\, \underline{x}(t_o) + \int_{t_o}^{t} e^{\underline{A}(t - \rho)}\, \underline{B}\, u(\rho)\, d\rho \qquad 6.1$$

Since we assume a sample and hold, the value of u is constant between sampling points.

Now, starting at any given t_o, the state of the controlled process is characterized by $\underline{x}(t_o)$ and control $u(t_o)$. At this point, the second fast -model is initialized, i.e.

$$\underline{x}'(t_o) = \underline{x}(t_o) \; ; \quad u'(t_o) = u(t_o) \qquad 6.2$$

The second fast model is released, and the first fast model is initialized. Since the first fast model is a feedback controller and does not contain any sampler, its control $e(\rho)$ is a continuous variable. Thus,

$$\underline{x}''(\beta) = e^{\underline{A}''(\beta - t_o)}\, \underline{x}(t_o) + \int_{t_o}^{\beta} e^{\underline{A}''(\beta - \rho)}\, \underline{B}\, K_m\, e(\rho)\, d\rho \qquad 6.3$$

$$t_o \leq \beta \leq t_o + T$$

From equation 4.30 we see that

$$u'(t_o + \mu) = \frac{E}{M}\, \underline{B}^T \int_{t_o + \mu}^{t_o + T} e^{\underline{A}'^T(\rho - t_o - \mu)}\, \underline{C}\, e_1''(\rho)\, d\rho \qquad 6.4$$

where

$$e_1''(\rho) = y - \underline{C}^T \underline{x}$$

The second fast model is forced in the interval $(t_o, t_o + \mu)$ by $u'(t_o)$. Hence,

$$\underline{x}'(t_o+\mu) = e^{\underline{A}'(\mu)}\, \underline{x}(t_o) + \int_{t_o}^{t_o+\mu} e^{\underline{A}'(\mu-\rho-t_o)}\, \underline{B}\, u(t_o)\, d\rho \qquad 6.5$$

Therefore,

$$\underline{x}'(t_o+2\mu) = e^{\underline{A}'(\mu)}\, \underline{x}'(t_o+\mu) + \int_{t_o+\mu}^{t_o+2\mu} e^{\underline{A}'(2\mu-\rho-t_o)}\, \underline{B}\, u(t_o+\mu)\, d\rho \qquad 6.6$$

or,

$$\underline{x}'(t_o+2\mu) = e^{\underline{A}'(2\mu)}\, \underline{x}(t_o) + e^{\underline{A}'\mu} \int_{t_o}^{t_o+\mu} e^{\underline{A}'(\mu-\rho)} \underline{B}\, u(t_o)\, d\rho + \int_{t_o+\mu}^{t_o+2\mu} e^{\underline{A}'(2\mu-\rho-t_o)}\, \underline{B}\, u(t_o+\mu)\, d\rho \qquad 6.7$$

At $t_o+\mu$, the first fast model is reinitialized and results in

$$\underline{x}''(\beta) = e^{\underline{A}''(\beta-t_o-\mu)}\, \underline{x}'(t_o+\mu) + \int_{t_o+\mu}^{\beta} e^{\underline{A}''(\beta-\rho)}\, K_m\, e(\rho)\, d\rho \qquad 6.8$$

giving an updated version of the error from $t_o+\mu$ to t_o+T. Call it e_2''. Hence,

$$u'(t_o+2\mu) = \frac{E}{M} B^T \int_{t_o+2\mu}^{t_o+T} e^{\underline{A}'^T(\rho-t_o-2\mu)}\, \underline{C}\, e_2''(\rho)\, d\rho \qquad 6.9$$

We can generalize these results,

$$\underline{x}'(t_o + n\mu) = e^{\underline{A}'(n\mu)}\,\underline{x}(t_o) + \sum_{n} \int_{t_o+(n-1)\mu}^{t_o+n\mu} e^{\underline{A}'(t_o+n\mu-\rho)}\,\underline{B}\,u(t_o+(n-1)\mu)\,d\rho \tag{6.10}$$

$$u'(t_o + n\mu) = \frac{E}{M} B^T \int_{t_o+n\mu}^{t_o+T} e^{\underline{A}'^T(\rho-t_o-n\mu)}\,C\,e_n''(\rho)\,d\rho \tag{6.11}$$

$$n = 1,2,\ldots,\frac{T}{\mu}$$

where

$$e_n'' = y - \underline{C}^T\underline{x}''(\rho) \tag{6.12}$$

with $\underline{x}''(\rho)$ the output of the first fast model over the interval $(t_o + n\mu, t_o + T)$.

With $\underline{x}'$ known (equation 6.10) in the interval (t_o, t_o+T), we can form a second estimate of the future system error e'. Hence, the control for the next sample period can be calculated. i.e.

$$u(t_o + \pi) = \frac{E}{M}\,\underline{B}^T \int e^{\underline{A}^T(\rho-t_o-\pi)}\,\underline{C}\,e'(\rho)\,d\rho \tag{6.13}$$

for any arbitrary t_o. Whenever real time t equals $t_o + n\pi$, the entire operation is started again, i.e. each fast model is initialized with the state variables of the plant at $t_o + n\pi$, and estimates of the future response over the new interval are made.

VII. SIMULATION RESULTS FROM TWO FAST MODEL APPROACH

Figure 7.1 shows a step response for the two fast model approach. It is immediately apparent that any differences of plant response from that of the one fast model approach, shown in figure 5.3, are quite subtle. However, a check of the output from the second fast model, or predicted output, was very different in this case. Consider the dotted curve starting at time 3. No longer did the predicted response dip toward zero as in figure 5.3, but generally followed the same pattern as the actual system output. As can be seen from the dotted curves, they predicted accurately for about 30 to 40 per cent of the preview span, then diverged from the system response.

A diagramatical explanation of this operation might be helpful. Consider figure 7.2 below.

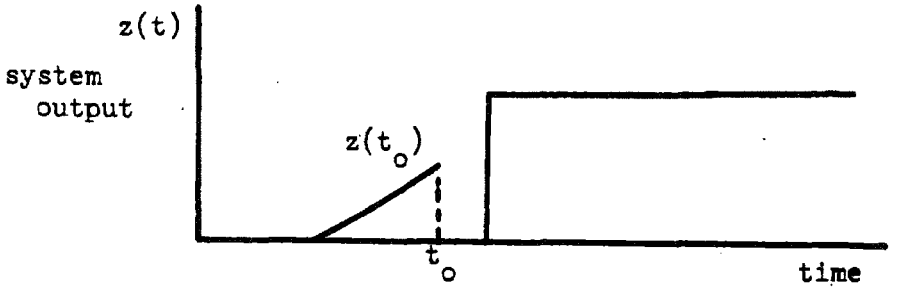


Figure 7.2

The curve indicates the process output up to time t_o. The curve between t_o and $t_o + \pi$ will be generated by the value of the control variable at t_o. The control at $t_o + \pi$ will be generated as follows. At t_o, the first model is initialized with $\underline{x}(t_o)$ and $u(t_o)$ as we have previously stated. Between t_o and $t_o + \mu$, the first model will produce an output something like figure 7.3.

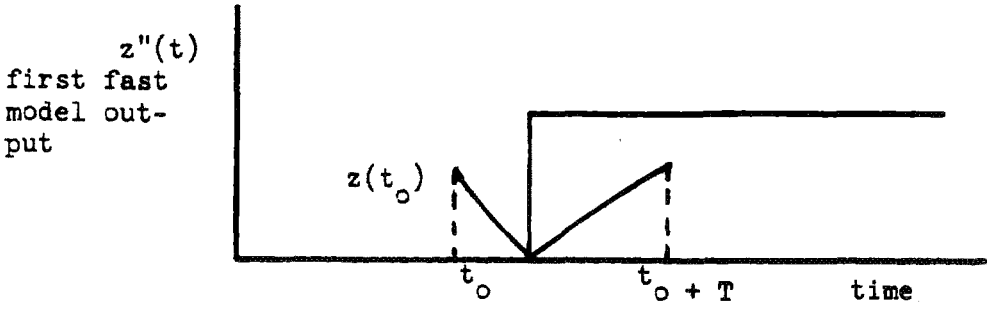


Figure 7.3

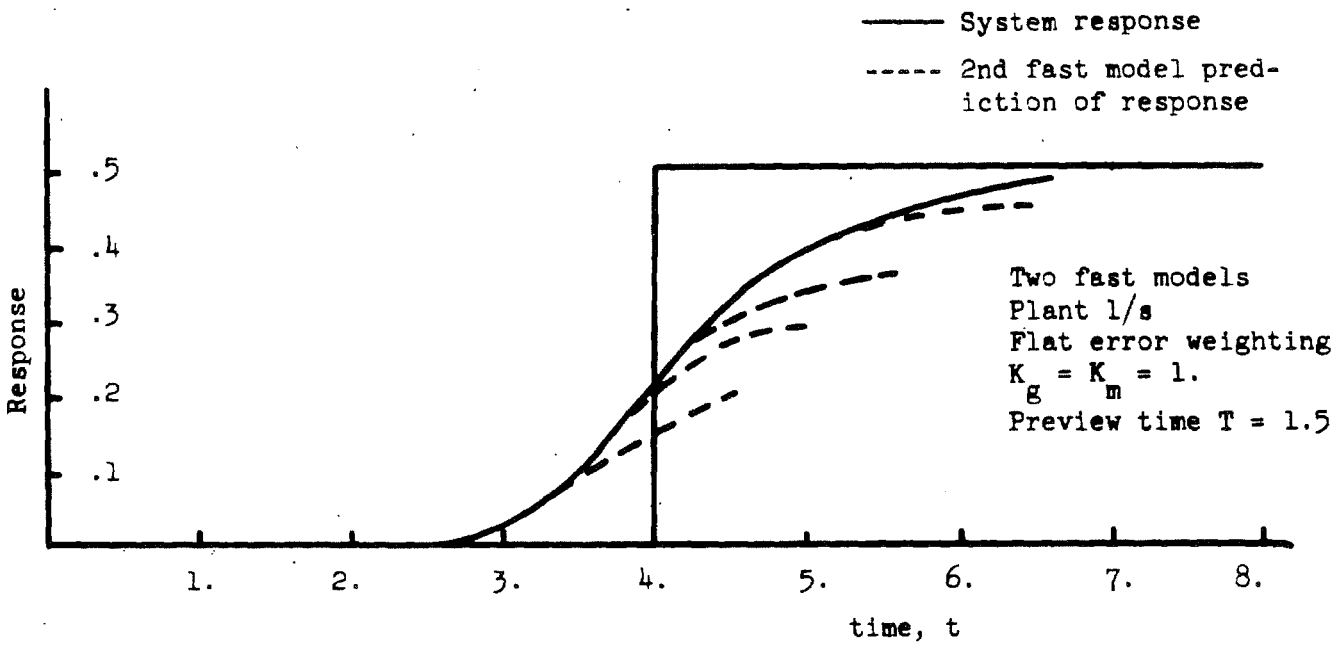

System response
2nd fast model prediction of response
Response
.5
.4
.3
.2
.1
1.
2.
3.
4.
5.
6.
7.
8.
time, t
Two fast models
Plant 1/s
Flat error weighting
$K_g = K_m = 1.$
Preview time T = 1.5

Figure 7.1

This will be weighted by the error weighting function to form the data from which $u'(t_o + \mu)$ is generated. At $t_o + \mu$ the first model is reinitialized and generates the following:

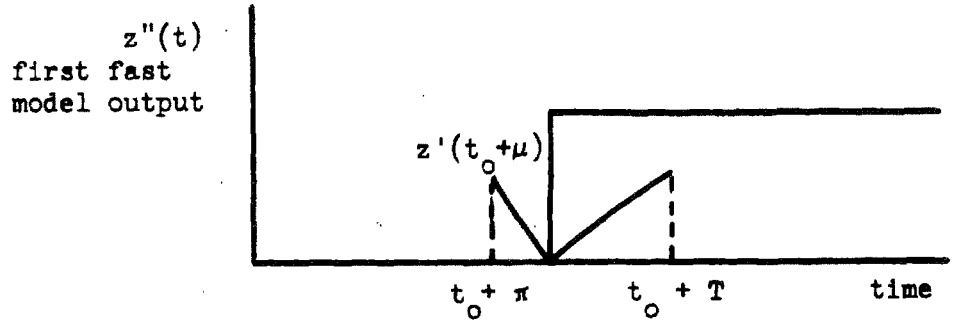


Figure 7.4

between $t_o + \mu$ and $t_o + 2\mu$. We then can determine $u'(t_o + 2\mu)$ and hence generate $\underline{x}'$ between $t_o + 2\mu$ and $t_o + 3\mu$.

At $t_o + n\mu$ the first model output will be something like the following:

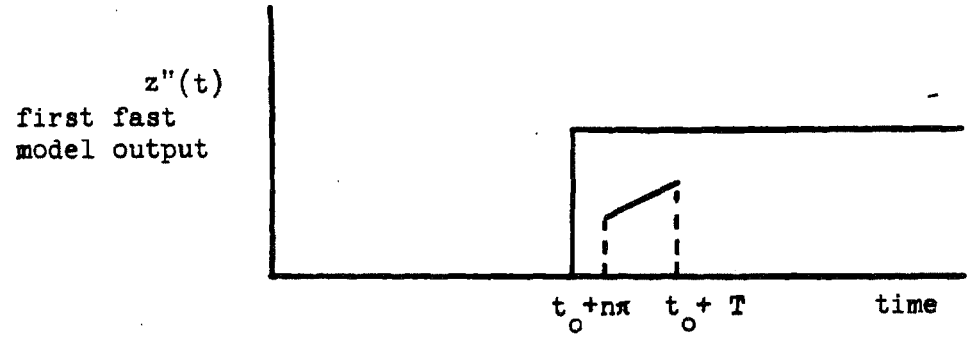


Figure 7.5

We can see that the first estimate of error is made over a constantly decreasing interval. After the complete sequence of first estimates is made, the second model output would look like the following:

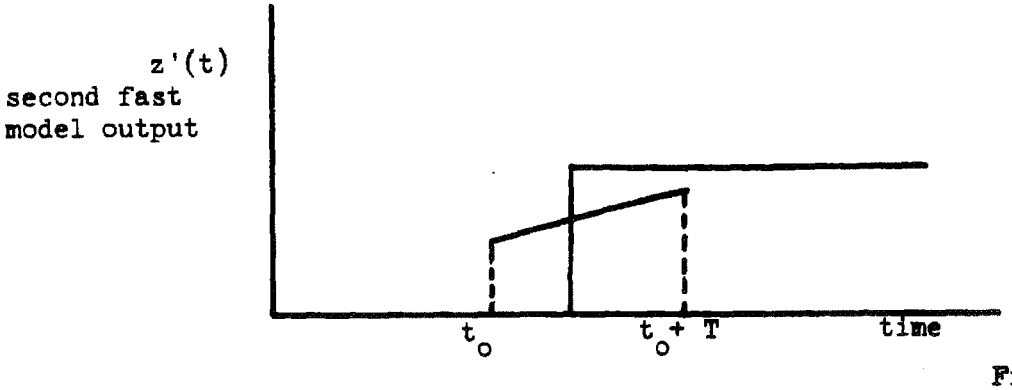


Figure 7.6

Hence, $u(t_o + \pi)$ can be obtained by weighting and integrating the second estimate of future error which is formed by comparing the output of fig. 7.6 with the known input. Then $\underline{x}$, the system state, is known in the interval $(t_o + \pi, t_o + 2\pi)$. Whenever real time reaches a sample point the models are reinitialized and the response over the next preview interval is again estimated. Notice that only the first error estimate is made on a shrinking interval. The second fast model which produces the second response estimate is reinitialized only at real time sample points, $t_o + n\pi$, and hence the available preview is always T. This means that if there are m sample periods π in a section of the given input, the second fast model executes m cycles, and the first fast model executes m^2 cycles of varying length. We have assumed that there are the same number of sample points in the scaled interval as in the real.

Appendix C contains step responses for a few other plants, and the corresponding second model output to indicate prediction quality. The integral type plants have the expected type of response, however, the proportional $1/(s + 1)$ plant is unusual. (see figure C-3). Since we were minimizing $\int (e^2 + u^2)d\rho$, and in the steady state, input equals output for this plant, the minimum occurs when the two are equal. This would not be practical for efficient control.

Notice also that the predicted response does not approach the actual system response as an asymptote. This at first glance appears to be an unusual result. However, if one writes the solution to the model response starting at any initial condition, such as at $t = 5$, one finds that the predicted response is correct.

The fact that this prediction is so poor simply reinforces the earlier statement that the control is not practical. The plant must be redesigned to provide better performance. The most obvious change would simply be increasing the gain, or other compensation such as the addition of an integrator in the forward loop could be used. If it is impossible to alter the plant, better performance in terms of error can be obtained by changing the performance index. Increasing E has the effect of raising loop gain, but at the expense of control.

The question of convergence to the true optimum value of the performance index H is important. For our computational method to be convergenent,

H must decrease with each successive approximation. That is, the magnitude of H must be smaller when two fast models are used than when one fast model is used, for convergence to be insured. H values for various plants are given in table 7.1.

For every case studied, the method was convergent. For well behaved plants the difference between the values of H for one and two fast models is small, but for less stable plants, such as $5/s(3s + 1)$, the level of iteration makes a significant improvement. Note, that in every case the addition of preview resulted in a very significant reduction in the performance index over the simple feedback control case.

PLANT	PERFORMANCE INDEX H			NORMALIZED INDEX H_n			PARAMETERS*	
	FEEDBACK	1 MODEL	2 MODELS	FEEDBACK	1 MODEL	2 MODELS	K_g	K_m
$\frac{1}{s}$	.2539	.1385	.1370	1.0	.547	.540	1.0	1.0
$\frac{1}{s(s+1)}$	.7766	.1818	.1732	1.0	.234	.223	1.0	1.0
$\frac{5}{s(3s+1)}$	50.7815	.6499	.1248	1.0	.0127	.0024	5.0	5.0
$\frac{5}{s(3s+1)}$	.5109	not checked	.0978	1.0	-----	.192	5.0	.24

* Error weighting function ω used was theoretical optimum in all cases. The values of H are minimum for the case stated.

TABLE 7.1

VIII. OPTIMUM PREVIEW SPAN AND SAMPLE PERIOD

We have made little mention of the preview distance used by the system. Since one of the objectives of the study was to determine if an optimum value did exist, this aspect had to be examined carefully.

In the derivation of the optimum control for the plant, preview time was assumed a given constant. However, experience and instinct would indicate that for the human operator some particular value of T would give a minimum performance index. If preview were completely unconstrained one would expect an infinite preview allowing maximum use of the information available.

To test this characteristic in our model of the human, various values of preview were used. Consider the data in figures C-2, C-5, C-9, and C-11. All plants studied showed the same general pattern with respect to preview, namely a distinct minimum value of performance index corresponding to a specific value of preview time.

These results were not surprising. Data by Bernotat and Widlok[7] in work with predictor displays showed similar behavior. Fogel[8] expressed the same type of idea in his stochastic model of the human in a control task such as driving an automobile or landing an aircraft. He presented a qualitative curve similar to those cited above.

Fogel's control strategy was similar to that expressed earlier in this thesis. His reasons for the unimodal shape of the curve were the following:

1. If the operator chooses to monitor points too far ahead of the vehicle, irrelevant information is introduced into the problem of determining the immediate control action.
2. If he monitors too close, the response time of the vehicle becomes a significant factor.

In Fogel's approach, noise consisting of wind gusts and effects of road roughness etc. were assumed to enter into the problem, as they certainly do. The operator performs some weighted average as a function of velocity and his independent estimates of the characteristics of the noise and vehicle response in determining his control action.

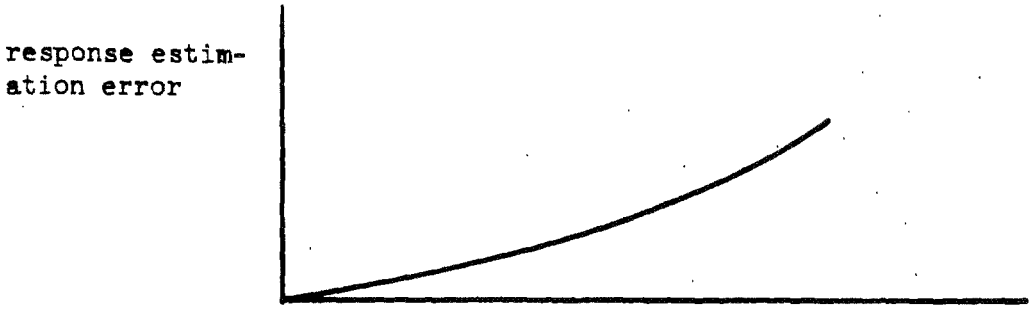
We were not concerned with noise aspects of the input signal. The input to the model was assumed to be perfectly known, or to be the result of some procedure for estimating the true signal from the noise. The ad-

ditional uncertainty remained however. Even with no noise, the operator cannot make a perfect error or response estimation over the preview span, unless the task is extremely trivial. This is due to his lack of complete knowledge of the dynamics of the vehicle and his limited memory and calculation ability. The more experience he has and the better his understanding of the plant, the better his response estimations will be. He has a better understanding of what will happen if he executes a certain control action, and he has a larger store of information upon which to base his control action.

The computer models we used have the same type of limitations. They cannot predict perfectly the future response. The more models used, or the more closely the iterations predict the actual response, the closer the plant response will approach the true optimum. However, the reasons for the differences between estimated response and the actual resulting response are probably quite different in the case of the human operator. We suspect the human does not calculate an adjoint or anything similar, yet somehow uses his experience to predict and based on this prediction does a reasonable job of controlling.

The computer has no experience and can only calculate. Since we have not given it perfect information, it cannot do a perfect job either. The more models we allow it to use in solving the problem, the closer to the optimum it will come.

Consider the following assuming a perfectly known input signal. The human might have a response estimation error versus preview time curve something like figure 8.1.

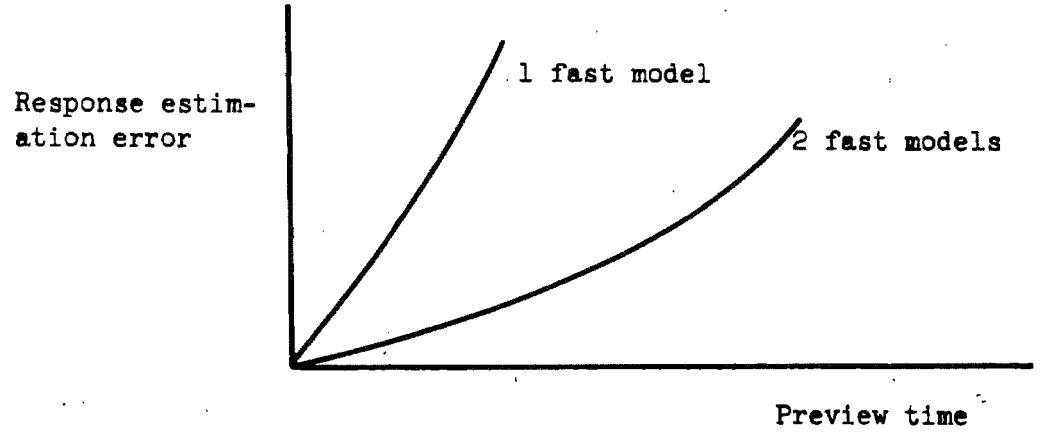


Presumed res. est. error vs. preview for humans

Figure 8.1

Response estimation error is defined as the deviation of the anticipated response from the actual response.

The computer model has a set of curves that look something like the following.



Res. est. error vs. preview for computer model

Figure 8.2

Note that simple feedback does not apply.

To justify the computer curves of figure 8.2, consider figures C-5,C-6, and C-7. Figure C-5 indicates that the optimum value of H for this plant occurred at T = 3.0, with a value of .212. With preview of 1.5, the accuracy of the predicted response was very good, but the performance index increased to .449. Hence, significant information was lost by using the shorter preview span. With preview of 4.5 units, the accuracy of the predicted response disintegrated greatly, and the system used highly incorrect information which resulted in a performance index H of .422. Hence, accuracy of the predicted response does decrease with preview and the shape of the curves in figure 8.2 is justified.

The reason the two fast model curve is lower than the one fast model curve is also increased accuracy. The additional fast model increases the prediction quality as demonstrated in section VII. Hence, the system can obtain significant information farther out in the preview span, with a corresponding decrease in the performance index.

This is especially apparent for under damped plants. Figure C-12 dramatically demonstrates this behavior. The additional fast model flattened the response estimation error curve indicated by the shift of

the curve toward higher values of preview. One would expect that including additional fast models would follow the same trend, i.e. increase the optimum span and decrease the performance index H. The effect of each additional model used should become less significant in terms of performance index. That is, as the true optimum is approached the change in H netted with each additional iteration will decrease.

As the human becomes a better controller, he should demonstrate the same type of behavior. Namely, his prediction error versus preview time curve should flatten, and he should be able to predict his response farther into the future.

Another assumption made in the formulation of the problem must be justified. We assumed the operator samples the surroundings and bases his control on this sample. First, the human cannot compute or decide upon a control action instantaneously so there must be a delay before control is executed. This corresponds to computation of the control by the fast models. Second, everyone knows from his own experience that the operator does not spend all his time staring straight down the road. He actually time-shares his senses--looking down the road, in the mirror, at the scenery, etc.

But how long should the sample period be? To check the variation of the performance index with the sample period , we slightly modified the regular programs. Instead of assuming a constant value for the control signal between samples, we allowed it to vary according to the equation for optimum control, equation 4.30. We updated the initial conditions on the models at the sample points as before, but between samples the control was made to follow equation 4.30.

The results for one fast model are shown in figure C-13. This plant was fairly undamped; hence results for better behaved plants will not be worse than this. Notice that for the case tested it made little difference if another sample was not taken until half of the preview span was traversed. However, when the sample period approached the preview period, the performance index increased by a factor of fourteen.

Back to the reason for discussing this parameter. By varying the sample period and the number of iterations used, the computer results can be manipulated to match the human characteristics in the same situation.

Another interesting point can be derived from figure C-13. If we used a sample period equal to the preview span, i.e. $\frac{\pi}{T} = 1$, exactly two iterations of the response over the interval $(t_o, t_o + T)$ were made, one by each fast model. However, by updating more often, i.e. $\frac{\pi}{T} < 1$, the resultant number of estimates was something greater than two. The equivalent number of iterations in the sense of the boundary condition iteration problem, where each guess covers the entire interval $(t_o, t_o + T)$, could not be determined due to the shrinking interval used by the first-fast time model to produce its estimate of the response.

By updating often, or keeping π small, a wrong estimate does not have time to seriously hurt the result. That is, the new scan of the input allows correction of poor control decisions before their result is harmful.

In terms of the human operator, updating allows him to obtain new information that might influence his previous action, provided he does not wait too long before updating.

IX. NON OPTIMUM WEIGHTING FUNCTIONS

All data to this point involved use of the error weighting function indicated by equation 4.31 for the particular plant under consideration. This equation was the result of theoretically minimizing

$$H = \int (Ee^2 + Mu^2)\, d\rho$$

We have not presented any data to validate the correctness of this result.

If one accepts the curves in figure 8.1 and 8.2 as correct, the first weighting function that comes to mind is of the form

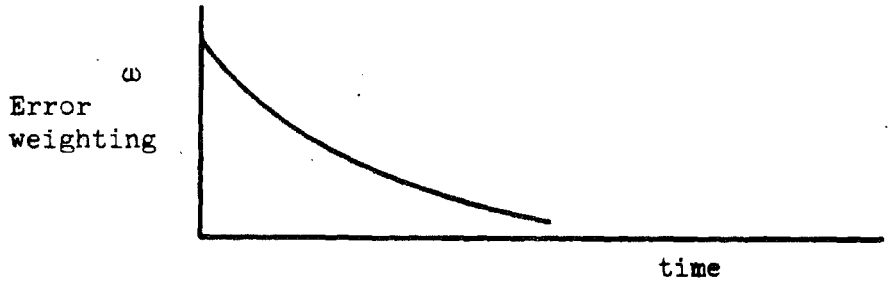


Figure 9.1

Since the accuracy of the prediction decreases as we move farther out into the preview span, one would expect some type of error weighting that is inversely proportional to the amount of error in the prediction. The reason for expecting this behavior comes from the definition of the error weighting function. It is an indication of the value placed on each point in the predicted interval, and since accuracy is decreasing with time, it is reasonable to expect less weighting of the inaccurate portions of the predicted response.

As an approximation to this approach a linear curve like that in figure 9.2 was used. A flat weighting function was also checked.

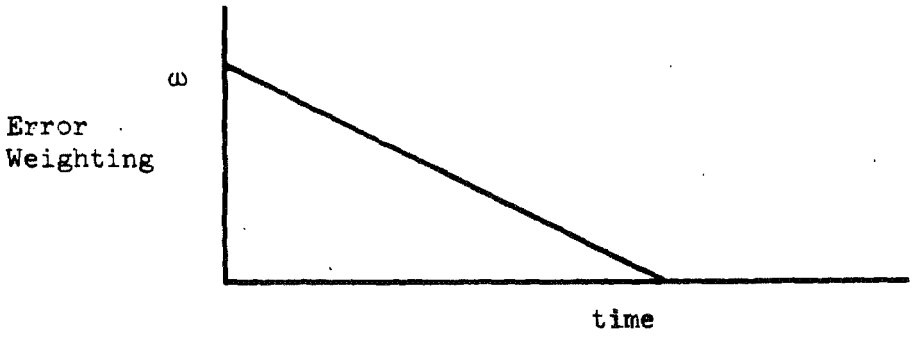


Figure 9.2

In no case did the theoretical weighting function of equation 4.31 fail. That is, no weighting function tried produced a lower performance index for a given plant than did the theoretical optimum function of equation 4.31.

Some very lnteresting results were obtained using these arbitrary weighting functions mentioned above, especially the decreasing function. With a simple $\frac{1}{s}$ plant, the performance index was quite good, and the trajectory shown in figure C-14 was quite similar to the response obtained using the optimum flat weighting function shown in figure 7.1. However, notice that the minimum performance index for the linearly decreasing weighting function occurred at a preview time nearly twice as great as the preview of 1.5 which produced the minimum H for the flat weighting function. A look at the prediction curves of C-14 and comparing them with the prediction curves in 7.1 shows the the quality of the decreasing function's predictions was not good,depreciating slightly faster than for the flat weighting function. The decreasing function forced the system to look farther into the future for information. This meant that it used highly inaccurate information, even though weighting it lightly.

For more complex plants, things were not so pretty. For example, consider figure C-16. This system was not well designed, but with the theoretical optimum weighting function it gave the reasonable results shown in figure C-10. Note that the response and performance index were quite good. For one fast model and a decreasing error weighting function the the predicted response was wildly unstable, but the final system response was nevertheless a good one with a very good performance index. If a human actually based his response on such an estimate, his reaction would undoubtedly be to stop as quickly as possible.

By lowering K_m, the gain of the first fast model, the predicted response shown in figure C-17 was better, but by no means acceptable. However, by virtually ignoring the prediction by weighting it lightly the overall system worked reasonably well.

The results from the two fast model configuration gave additional relevant information concerning the overall system performance. As shown in figure C-18, the performance index was lower for one fast model than it

was for two fast models when using the linearly decreasing weighting function. This meant that the calculation procedure was not convergent for this type weighting function. That is, h did not decrease with successive iterations. Hence, considerable difficulty could be encountered by using this weighting function.

Notice that the theoretical optimum weighting function provided a lower performance index than either one or two fast models using the decreasing weighting function. (see figure C-18)

These results show the way in which the available preview should be used. For a given plant and criterion for judging performance, there exists an optimum weighting function described by equation 4.31. Inherent in the system there is a specific preview time, depending on the various inaccuracies present in the system, which minimizes the performance index for that given system. This means that human operators should not necessarily use all available preview, but only that amount which gives sufficient information to minimize the performance index.

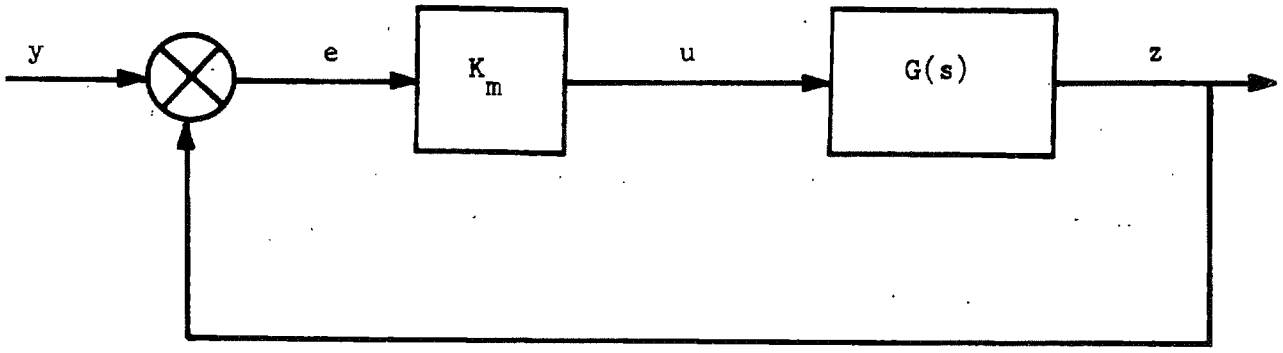
Best results are obtained when the anticipated response is treated as absolutely true, even though it is known to be slightly in error, and weighting with the optimum weighting function. Updating information frequently nullifies the effect of innaccurate prediction.

X. DESIGN OF THE FIRST FAST MODEL

The configuration of the first fast model in the three-time-scale system or the feedback model in the two-time-scale system is not determined by the optimum control equation, but is an extremely important part of the complete system. Since this model is essentially the first iteration, it determines to some degree the speed of convergence to the optimum value of H.

The simplest method is to use the multiplicative factor from the optimum equation as the value of K_m. In fact, most of the data was taken this way and the results were very satisfactory in most cases.

The next more complex method would be to optimize the value of gain subject to the performance index H given the general configuration of the loop, figure 10.1.



Simplest configuration for first fast model

Figure 10.1

This is still a simple procedure and results have been tabulated in many references. The general procedure is outlined in chapter 2 of reference 9.

For example, for $G(s) = \frac{5}{s(3s + 1)}$ and performance index of the form

$$H = \int (e^2 + u^2)\, d\rho$$

and

$$u = K_m\, e(\rho)$$

the best value of K_m was calculated to be approximately 0.3.

To prove this result, the following data was taken. (see figure C-19) Using no preview, hence simple feedback control, the value of K_m was varied and the resulting performance index compared. The best performance occurred at K_m = .24. The difference between the simulated value for optimum K_m and the calculated value id due to numerical inaccuracies in both the simulation and the calculation procedure. However, such a small difference in K_m was virtually indistinguishable when the feedback system incorporating K_m was used as the first fast model in the three time scale system. Consider the table below.

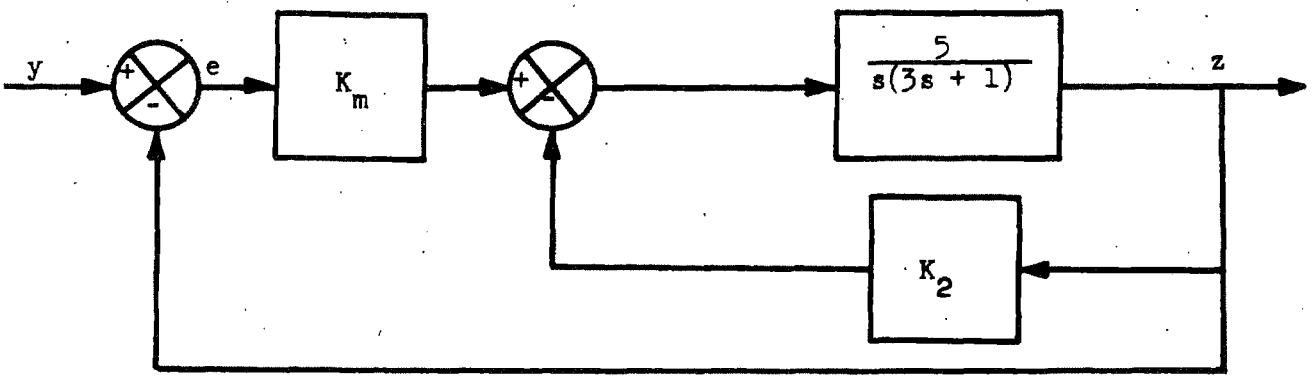
PREVIEW	$\int u^2$	$\int e^2$	$\int e^2 + u^2$
K_g = 5.0 K_m = .29			
1.4	.039	.070	.109
1.6	.028	.069	.097
1.8	.020	.104	.124
K_g = 5. K_m = .24			
1.4	.039	.070	.109
1.6	.028	.069	.097
1.8	.019	.108	.127
K_g = 5. K_m = 5.			
1.4	.070	.055	.125
1.6	.119	.033	.152
1.8	.151	.027	.178

For gains in the range of K_m = 0.3 there was no difference to three decimal places.

However, the improvement in performance over the arbitrarily selected value K_m = 5. is very significant. Notice that the higher gain system had a lower error as might be expected, but the control was three times the magnitude of the control when a lower gain was used, and the performance index increased by 22 percent.

The effect of this change in gain is best shown by figure C-20.

By changing the gain we have decreased the amplitude of the oscillations, but have not changed the rate of decay. From this diagram one might think that compensation, i.e. a rearrangement of the poles, might be a logical method to improve performance. A system of the form shown in figure 10.2 with $K_2 = .3$ was tried.



Compensated fast model configuration Figure 10.2

Optimum K_m for this arrangement was .50. This system reduced the performance index for the simple feedback case to .356 compared to .511 for the uncompensated system. Notice the compensation used was essentially rate feedback, and would seem like a logical method since the process is receiving preview information and derivative information should be a good first approximation.

However, when this system was used as the fast-time predictor model, the performance index did not improve. For two fast-time models using the optimum weighting function, the performance index was .0982 for the compensated first model, and .0978 for the uncompensated model.

These results are not extensive enough to be considered conclusive by any means, but give important hints at the design procedure for such a preview system. The use of a compensated first model does not seem necessary, however optimizing the parameters of the first model is important. Notice, that the performance index with the first fast-time model 20 times optimum still gave a result better than any non-preview case studied. This would indicate that the exact configuration is not critical unless performance index requirements are very tight.

There is another possible deviation in the configuration of the first fast model. It was pointed out earlier that the estimated response from this model is made over a shorter and shorter span (see figures 7.4 and 7.5) due to the finite preview of the system. It was thought that assumptions allowing extrapolating the input farther into the future might be useful.

To check this we assumed that the input did not change from the value at the end of the preview span. The results were definately worse than the normal procedure of using the decreasing interval. Every case utilizing this assumption that we tested resulted in performance index increases of about 50 percent over the same case making no assumptions. Combining this with earlier results, all indications imply that it is best to make no assumptions about the first model, but simply use the dynamic model of the plant in a simple feedback loop.

XI. PERFORMANCE WITH CONTINUOUS INPUTS

Based on the step response, the preview time which minimizes the performance index is optimum for any given speed or rate of forced pacing. Therefore, due to the linear relation between velocity and time, an operator should always look a fixed time down the road, with the distance he looks varying according to the velocity he is traveling. Hence, when the input is too rapid or if high frequency components are present, the velocity should be decreased. This means the data is taken from a shorter interval in the same amount of time allowing the system to keep up with the input.

Since the present system is forced paced, the above cannot be executed automatically. The same result can be shown by decreasing the frequency of the input.

To demonstrate these characteristics, an input consisting of two sine waves was used. In figure C-21 the input was beyond the dynamic capabilities of the system. Due to the low pass characteristics of the closed loop system, the high frequency component was filtered out, and the fundamental passed slightly attenuated with some phase lag. No amount of preview could help this situation when restricted to a forced paced system. The obvious control strategy is to slow down the rate at which the input is presented to the system.

To model such a strategy, the input was changed to

$$y = .5(\sin .2t + .4\sin .8t)$$

and the results are shown in figure C-22. Since distance = velocity x time, we reduced the distance by a factor of five, while maintaining a constant preview span. In this case, the preview controller did a very good job. It followed with a slight phase lag , but the difficulties of the previous trial were largely eliminated.

When the input was slowed even more to

$$y = .5(\sin .1t + .4\sin .4t)$$

the response became near perfect. In driving we do exactly the same thing, reducing speed when the road becomes dangerous in order to bring the quality of our control within satisfactory limits.

For a comparison between straight feedback control and preview control

with continuous inputs consider figure C-24. For the slowest input the feedback controller did follow the input, but with a high frequency oscillation about the signal. The comparative indices are revealing.

For preview (two fast models, optimum weighting, T = 1.6)

$$\int \text{error}^2 = .028$$
$$\int \text{control}^2 = .017$$
$$\text{sum} = .045$$

For no preview with optimum gain

$$\int \text{error}^2 = .085$$
$$\int \text{control}^2 = 2.139$$
$$\text{sum} = 2.224$$

there was a considerable difference in error, but the big difference was in the control effort used. Preview virtually eliminated control effort indicating that with a reasonabley smooth input an operator with preview can follow with very little effort on his part and keep errors to a minimum.

XII. DISCONTINUOUS IMPORTANCE WEIGHTING FUNCTIONS

One of the features discussed in the development of this modeling technique was non-uniform importance of the input. In some situations errors may not be very important at certain sections of the input, while at others, such as bridges, or parked cars, the error tolerances might be very tight.

To test this property in the model, the input shown in figure C-25 with importance weighting as indicated was used. Figure C-25 also shows the results for a plant with transfer function $\frac{5}{s(3s + 1)}$ using the optimum weighting function and preview of 1.6. Several difficulties can be noted. Since the system was second order, it could not change position instantaneously, and since it was highly underdamped, once started it tended to overshoot. Also, due to limited preview, it could not prepare for the step far enough in advance to utilize the low importance areas. When it did see the step, the high importance required rapid elimination of large errors causing the excessive overshoot. Compare this result with figure C-26 where uniform importance was used. The start was essentially the same, but overshoot was lower and the ramp cleanly executed.

Many of the above problems were due to the relatively unstable plant. The results for a single integrator shown in figure C-27 are prettier, but have the same basic limitations. Because of the limited preview, the low importance areas were not utilized.

Now, let us consider what we have done. The systems were set up with optimum parameters based on continuous uniform importance and no longer apply. Looking at the equations again, we still desire to minimize equation 4.4, but the value used for error is not the same.

$$e' = Ie \qquad 12.1$$

where I is the importance. Therefore H becomes

$$H = \int_0^T (EI^2e + Mu^2)\, d\rho \qquad 12.2$$

The effect of the importance weighting obviously is a change in the value of E, and will not change any of the variations since it is a

known input. The control then becomes

$$u(t) = \frac{B^T}{M} \int_{t}^{t_o + T} e^{A^T(\rho - t)} \underline{C}\, I^2\, e(\rho)\, d\rho \qquad 12.3$$

The optimum multiplicative gain is now I^2E instead of simply E. The use of importance other than uniform changes the gain of the system, which can cause stability problems and cause changes in the optimum value of other parameters, especially the preview span T.

Consider the curves shown in figure C-28. A value of $K_g = 2$ meant we were trying to minimize $\int 2e^2 + u^2$ while $K_g = 4$ meant we were minimizing $\int 4e^2 + u^2$. The data indicates that the higher the importance on error, the shorter the preview should be for optimum performance.

Recall the data for plants of the form $\frac{D}{s(\beta s + 1)}$.

Rewriting the above we see that the total multiplicative gain for the system is $K_g \times \frac{D}{\beta}$. That is, K_g from the control equation, and $\frac{D}{\beta}$ from rewriting the equation for the transfer function as

$$G(s) = \frac{D/\beta}{s^2 + (1/\beta)s}.$$

Hence, for $K_g = 1.$, and $D = 1$, $\beta = 2$, the gain was .5. With this system the optimum preview was 3.0 as shown in figure C-4. With $K_g = 1.$, $D = 1$, and $\beta = 1.$, the optimum preview was 2.5 units, and the gain obviously was 1.0. For $D = 5$, $\beta = 3$, and $K_g = 5$, the gain was 25/3 and the preview span for optimum performance dropped to 1.6 units. Although the time constants changed the damping in the cases listed above, the dominant factor in determining optimum preview was the overall gain of the system. As gain increased, preview time decreased.

To further verify this trend, we tried to minimize

$$\int (10e^2 + u^2)\, d\rho$$

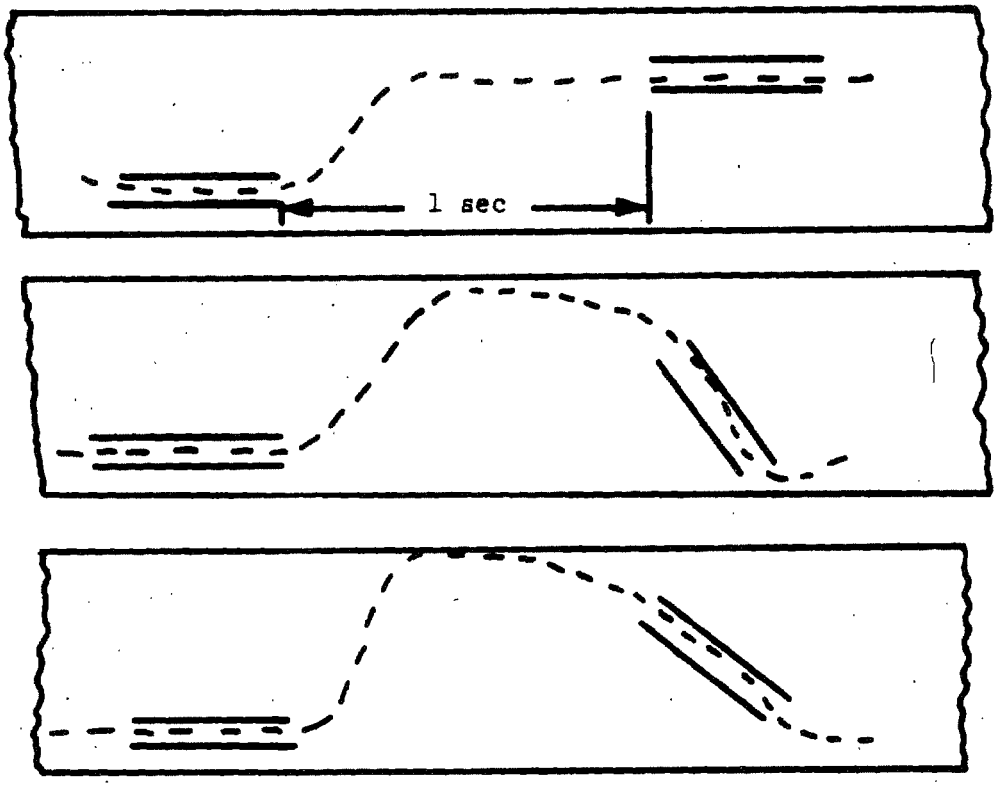
With preview of 1.2 or greater, the output was unstable; recall that for equal weighting of error and control, optimum preview was 1.6 units. The

exact parameters for minimizing $\int (10e^2 + u^2)\, d\rho$ were not found, but the preview for minimum H was less than 1.0 units.

The above indicates the reason for increased complexity in non-uniform inportance situations. The performance index changes with time requiring the preview to change if good performance is to result. For example, when the importance is low, the system should look far down the input path in order to prepare for future areas of tight tolerance and yet use a minimum amount of control. When regions of high importance are reached, less preview should be used in order to concentrate on the immediate area, effectively tightening the loop at these points.

This is not to be confused with the problem described in the previous sectionconcerning self-pacing. There, one preview time was used, varying the preview distance. Here we are actually changing the preview time. Obviously, if it were possible, the system would slow down, and use less preview in order to obtain the maximum amount of information from the surroundings in tight tolerance areas.

This is verified by the human subject data shown in figure 12.1. This subject was not particularly concerned about minimizing control effort, but he essentially performs as described above. After passing one obstacle and barely seeing another one far down the path, he moves the pen as far to the left as possible in order to prepare for the next bridge. This data was taken using the same apparatus that is described in the next section, except the subject had a perspective view of the input path. The figures are taken from reference 10.



Human subject data with discontinuous inputs
and non-uniform importance

Figure 12.1

XIII. HUMAN SUBJECT DATA

In order to validate this modeling technique a limited amount of data was taken using human subjects performing a simple preview control task. The apparatus used is shown in figure 13.1. The subject controlled the horizontal movement of the oscillograph pen by turning a steering wheel which was connected to a potentiometer. Analog computer dynamics were placed between the control wheel and the pen to simulate the plant. The subjects were given an unconstrained preview of the input which approached them at a constant rate. That is, they were allowed to use any amount of preview up to a maximum of about 3 seconds.

Two subjects were used. Subject one was the author, subject two was a female with no technical background. They were not required to minimize any specific performance index, but were told to follow the input path using any control strategy desired. They practiced enough to become reasonably familiar with the dynamic behaviour of the system, but were not given extensive training.

Two plants were used, namely $\frac{1}{s}$, and $\frac{1}{s(s+1)}$. Several step responses were generated by each subject for each plant. An average response was calculated for each subject-plant combination and the average curve was matched by a similar step response from the preview control system. Each input consisted of a succession of ramps and steps of various sizes and directions, with data taken only from steps .5 units in height.

For the single integrator, good results were obtained. In both figure C-29 and C-30 it was possible to match quite closely. The preview control system trajectory fell within the data spread of the human response. The same was true for $\frac{1}{s(s+1)}$ as shown in figures C-31 and C-32.

Some interesting conclusions can be drawn from this data. First, in the integrator case, both subjects were matched by the preview system using two fast models. In the more complex case they were matched by the system using one fast model. Since the dynamic behavior of the integrator is straightforward, its characteristics were quickly learned by the subjects. Hence, they were reasonably sure of its response given a certain input. Their knowledge of the more complex plant was not as good and their performance was modeled by a poorer quality model.

Both subjects used higher gains and shorter preview with the integrator, and both demonstrated a saturation effect. This was not saturation of the equipment, but of the subjects. They simply did not turn the control wheel except when starting and ending lateral motions. This hints that they were weighting changes in control more heavily than the magnitude of the control signal. This seems quite logical since no effort on the part of the subjects was required to hold the wheel in a fixed position away from the null. The fact that the saturation was lower for the female might have been the result of differences in physical strength, but more likely it was due to her difficulty in controlling overshoot. The data spread of Figure C-30 shows that she did tend to overshoot. By saturating at a low value, she was able to prevent greater difficulty.

With the more sluggish plant, she did not show a saturation effect, as can be seen from figure C-32. However, the response was not as fast, as with the integrator indicating that control levels were below her previous saturation level. She also used lower gain and longer preview implying that she was less concerned about error, but more concerned about control.

The author demonstrated a saturation effect with both plants. The reason that this saturation was lower for the second order system is not clear. It appears that changes in control were weighted very heavily. A longer preview span and lower gain was used with the $\frac{1}{s(s+1)}$ than with the integrator. All these seem to imply concern over control effort.

Further indication that neither subject understood the behavior of the second order plant is the flat error weighting function. Use of this function means that errors close to the present time were weighted equally with future errors. Because of the slow response of the plant these errors cannot be eliminated quickly. If this had been realized, a weighting function that places higher value on future errors probably would have been used.

Both subjects tended to weight error more heavily with the quick responding plant and weighted control more heavily with the sluggish plant. They used less preview when error was dominant, and both weighted change

in control quite heavily. All except change in control, which was not discussed, agree with results presented in earlier sections.

It would be difficult to state the exact criterion used by the subjects in performing this task, but nevertheless, their responses were modeled very well by the preview control systems discussed in this thesis and their general behavior explained in terms of the parameters of these models.

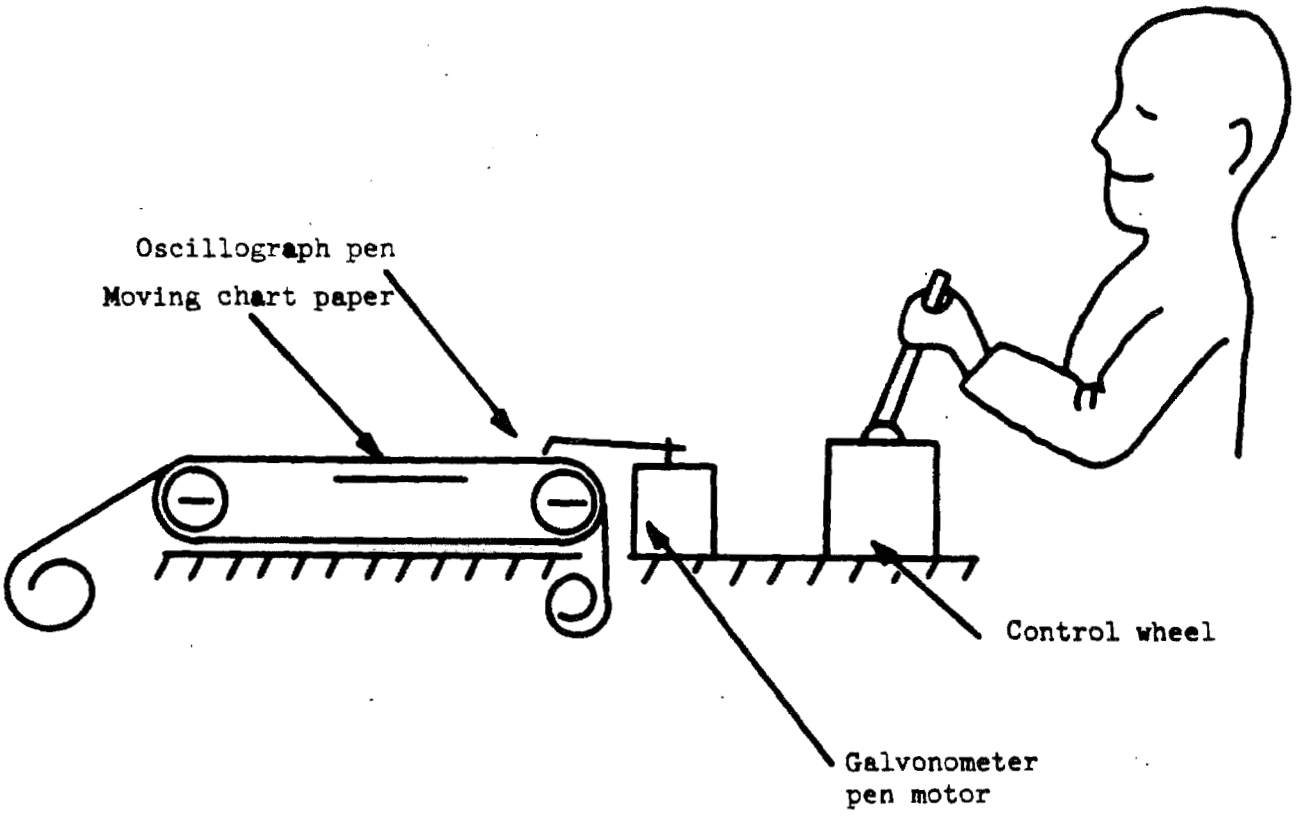


Figure 13.1 Apparatus for human subject data

XIV CONCLUSIONS AND RECOMMENDATIONS

I.) The first and most general conclusion is that the preview control modeling technique developed here does provide a reasonable model of creatures with the ability to look ahead at future inputs.

2.) We also see that the advantages of preview become more significant the more unstable (relatively speaking) the plant is. For simple, well behaved plants preview provides a lower preformance index, but the improvement is not as significant as that for a more complex plant. Also, the second fast model does not provide much improvement unless the plant is at least second order. The faster the response, and the less stable the plant, the greater the advantage of the second fast model. Hence, significant decreases in performance index are especially apparent in high gain, under damped systems.

We should expect this better performance index improvement with less stable plants, since it has been known for some time that the predictor instrument makes it possible for the human operator to easily control plants with two, three, or more integrations. While feedback, error nulling control works well with stable plants, the human has considerable difficulty using instantaneous error nulling techniques in controlling less stable plants. With these under damped plants, preview has the affect of additional damping, slowing down high frequency oscillations.

3.) We also show that there exists an optimum preview for any given system, depending on the controlled plant, importance, and accuracy of the fast-time predictor model. Generally, for the same class of plant, the higher the gain and quicker the response, the shorter the preview should be. For sluggish plants, longer preview is needed to allow sufficient time to prepare for future changes in the input. Changes in importance are equivalent to changes in gain.

No hard rule was found for the magnitude of the preview span, but in general it is a trade-off between accuracy and speed of re-

sponse. If the preview is too short, accuracy of the predicted response is very good, but significant information is not used. If too much preview is used, the degraded accuracy netted an increased performance index. The best performance is somewhere in the middle, and as long as the updating of the model is fast enough, that is, sampling period short enough, the slight innaccuracy in the predicted response causes no significant harm.

4.) We also show that given a fixed plant, the best results occur when no compensation is included, but existing parameters are optimized. Of course, if the plants are not fixed better results can be obtained through compenastion of various types.

5.) Another important point is the degree of accuracy required in modeling the actual plant for the predictor models. This turns out to be a matter of the degree of performance required. Even poor quality preview and prediction results in significant changes in performance index. It would probably be a matter of economics in determining how good you make the model. For mechanization of a preview control system and a well behaved plant, one fast-time model would probably be all that could be justified, in terms of performance index, unless more accurate predictions are required. Marginally stable systems would require two or more fast models.

6.) It is important to note that with the optimum weighting function and small sample periods, the preview model will not be more unstable than the plant itself. Thus, if the lowest order fast model, i.e., the feedback controller, is stable, the entire system will be stable provided the preview span is near optimum.

7.) If a performance index is known, the form of the best weighting function can be determined. For analytical situations this is quite simple, but in real situations this can be a significant problem. It should also be remembered that arbitrary weighting functions, even when providing reasonable performance indices, may lead to some disturbing

situations, namely that successive interations of response may not be converging to the true optimum. This was very apparent in section IX.

8.) Finally, the main reason for studing the preview model is to determine why preview is useful. It is obvious that preview allows antieipation of future transient, but preview is more important than that. We see that any optimization problem requires knowledge about the future response, thus becoming a preview problem. The human in a preview situation is one of a privileged few systems which have both a knowledge of future inputs and a relatively good estimate of what response will occur. He therefore has certain of the tools needed for solving the optimization problem. His response will be based essentially on experience and knowledge of the plant he is controlling, but still involves an estimate of future error. His response probably will not be truly optimum, but will be much better than any non-preview controller. Also, by using a sampled approach and updating before proceeding very far down the path, he has the opportunity to view the results of a bad decision before it causes significant damage. This means he has the potential to return to optimum control after only a short distance. This sampling also allows him to continually use optimum control over a long interval, even when he has only finite preview. Most optimization schemes require knowledge of the input all the way to the end.

Thus, the human, when combining esperience and preview, can approximate the optimum for a given control situation. This is an oversimplification, but it is a reasonable explanation for the value of preview in a control situation.

RECOMMENDATIONS

1.) Earlier restrictions such as forced-pacing and fixed preview were pointed out. Future work should include provisions for more realistic methods of controlling both velocity and preview.

2.) Only a bare minimum of experimental work with subjects has been done. More extensive work should be done to determine if a more fine

correlation between the model and human subjects does exist as expected. More and better trained subjects should be used.

3.) There may also be similarities with the predictor displays of the Kelley type. Although the predictor display is a different and more complex experimental task, it might be possible to determine why these instruments are so successful.

4.) Also, studies with noisy signals and constrained parameters, such as saturation of the controller, should be done.

There are numerous options available for study in this area. Based on the results obtained from the computer simulations, this should be an area of fruitful experimental endeavor.

REFERENCES

1. Zeibolz, H., and Paynter, H. M., "Possibilities for a Two-Time Scale Computing System for Control and Simulation of Dynamic Systems," Proc. National Electronics Conf., Vol. 9, pp. 215-223, 1953.

2. Kelley C. R., "Developing and Testing the Effectiveness of the Predictor Instrument," Office of Naval Research Tech. Rept. 252-60-1, Dunlap and Associates, Inc., Stamford, Conn., 1960.

3. Sheridan, T. B., "Three Models of Preview Control," IEEE Transactions on Human Factors in Electronics," Vol. HFE-7 no. 2, pp. 91-102, June 1966.

4. Johnson, W. M., "Two Time-Scale Predictor Model Using Scan of Anticipated Input," S. M. Thesis, Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, April, 1965.

5. DeRusso, Roy, and Close, State Variables for Engineers, John Wiley and son, New York, 1966.

6. Hildebrand, F. B., Methods of Applied Mathematics, Prentice Hall, Englewood Cliffs, N. J., second edition, 1965.

7. Bernotat, R., and Widlok, H., "Prediction Display," Institute for Guidance and Control, Technical University of Berlin, presented at the Tri-National Symposium on Information Presentation, Bedford, Mass., October, 1965.

8. Fogel, L. J., Biotechnology: Concepts and Applications, Prentice Hall, Englewood Cliffs, N. J., 1963.

9. Merriam, C. W. III, Optimization Theory and the Design of Feedback Control Systems, McGraw Hill, New York, 1964.

10. Sheridan, T. B., Johnson, W. M., Bell, A. C., Kreifeldt, J. G., "Control Models of Creatures Which Look Ahead," Proceedings of the Fifth National Symposium on Human Factors in Electronics," May 5-6, 1964, San Diego, California.

APPENDIX A BRIEF DESCRIPTION OF THE KELLEY PREDICTOR INSTRUMENT

The predictor instrument is designed to give the operator of a control system information about the future of the variable he is controlling. This prediction of the future response is based on the present state of the process and some assumption about the forcings. A block diagram of this system is shown in figure A-1.

A fast-time analog of the controlled plant is repetitively clamped with the state variables of the system at time T. The model is released, and quickly runs its course forming a prediction of the future response. This is repeated several times a second. This prediction is displayed on an oscilloscope. Hence, the operator can view his present position at time T, the present and also view the predicted response of the plant.

The similarities between this operation and the preview control model shown in figure 2.2 are obvious. The aim of the model however is to eliminate the human from the loop, describing his control action in terms of weighted averages.

There is a difference in this predictor display type instrument and the other type preview control model discussed in figures 2.3 and 6.1. These models assume the operator does not hold a constant value of control over the prediction interval but continuously varies control according to some criterion of performance. This requires a more complex fast time model.

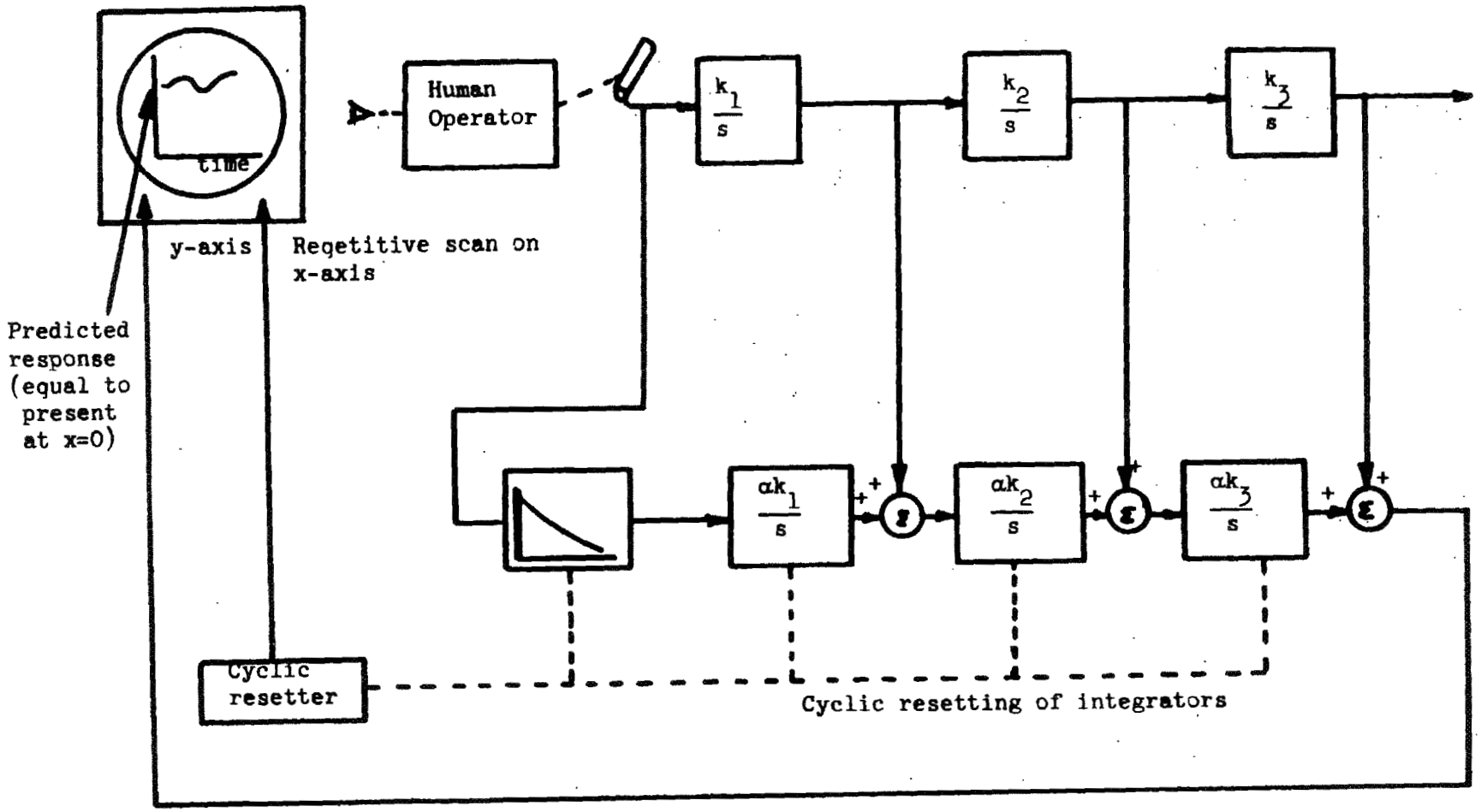


FIGURE A-1 KELLEY'S PREDICTOR DISPLAY FOR HUMAN OPERATORS

APPENDIX B COMPUTER PROGRAM DESCRIPTIONS AND FLOW CHARTS

The basic approach to the programming problem was the following. A main calling and computational program which utilized several external subroutines for calculating inputs, integrals, weighting functions and transition matrices was written. This allowed maximum flexibility in simulating any given situation. Since the main calling program evolved with the existing need, no claim for its efficiency can be made. It worked and provided the necessary data is all that can be said for it.

The dynamic systems were programmed using a state space approach. The solution to the first order differential equation of motion for a time invariant system is

$$\underline{x}(t) = e^{\underline{A}(t)}\,\underline{x}(0) + \int_0^t e^{\underline{A}(t-\rho)}\,\underline{B}\,u(\rho)\,d\rho$$

If we are interested with results only at discrete intervals and assume the control constant over this short interval, the equations can be written

$$\underline{x}(n\pi + \pi) = e^{\underline{A}\pi}\underline{x}(n\pi) + \int_{n\pi}^{n\pi+\pi} e^{\underline{A}(n\pi+\pi-\gamma)}\,\underline{B}\,d\gamma\;u(n\pi)$$

where π is the length of the interval.
With a change of variables,

$$\mu = n\pi + \pi - \gamma$$

$$\underline{x}(n\pi) = e^{\underline{A}\pi}\underline{x}(n\pi) + \int_0^T e^{\underline{A}\mu}\,\underline{B}\,d\mu\;\;u(n\pi)$$

For a given period π, both the matrix exponential and the integral above are constants easily programmed for solution. As long as π is small, the results are very accurate.

The use of a subroutine to calculate the matrix exponential was not necessary for simple plants, but was included for versatility and convenience. Provisions were also included for simulating saturation effects in the controller. No studies were done using this facility, but it was necessary for matching human data. The subroutine for calculating integrals was a version of Simpson's rule for numerical integration.

Flow charts for the main program are presented at the end of this appendix.

An important characteristic of any computer program is its running time. The biggest difficulty in this area comes when using two fast models and long preview spans. For each data point in real time the following must be calculated:

$$A = \frac{\text{length of preview span } T}{\text{length of sample period } \pi} \quad \text{points in the second fast mod.}$$

For each point calculated in the second fast model

$$\frac{\text{Preview span } T - \text{present time}}{\text{sample period } \pi} \quad \text{points calculated by}$$

the first fast model. Hence

$$B = \sum_{0}^{\text{PREV.}} \frac{\text{preview span} - n}{\text{sample}} \quad \text{points must be calculated by the first}$$

fast model for each data point in real time.

This means A + B points must be calculated for each point in real time. Each calculation is made by passing through a short loop in the computer program. In addition to the above, auxilary information such as performance indices and output must be generated. Calculation time can be kept to a minimum by using shaort data runs. If π is lengthened, the running time decreases, but so does the accuracy.

For example, if a data set 15 units long with preview of 2 units and sample period of .1 units,

$$15*\left(\frac{2.0}{.1} + \sum_{0}^{20} \frac{2.0 - n}{.1}\right)$$

$$= 15*(20 + 20 + 19 + \ldots + 1)$$

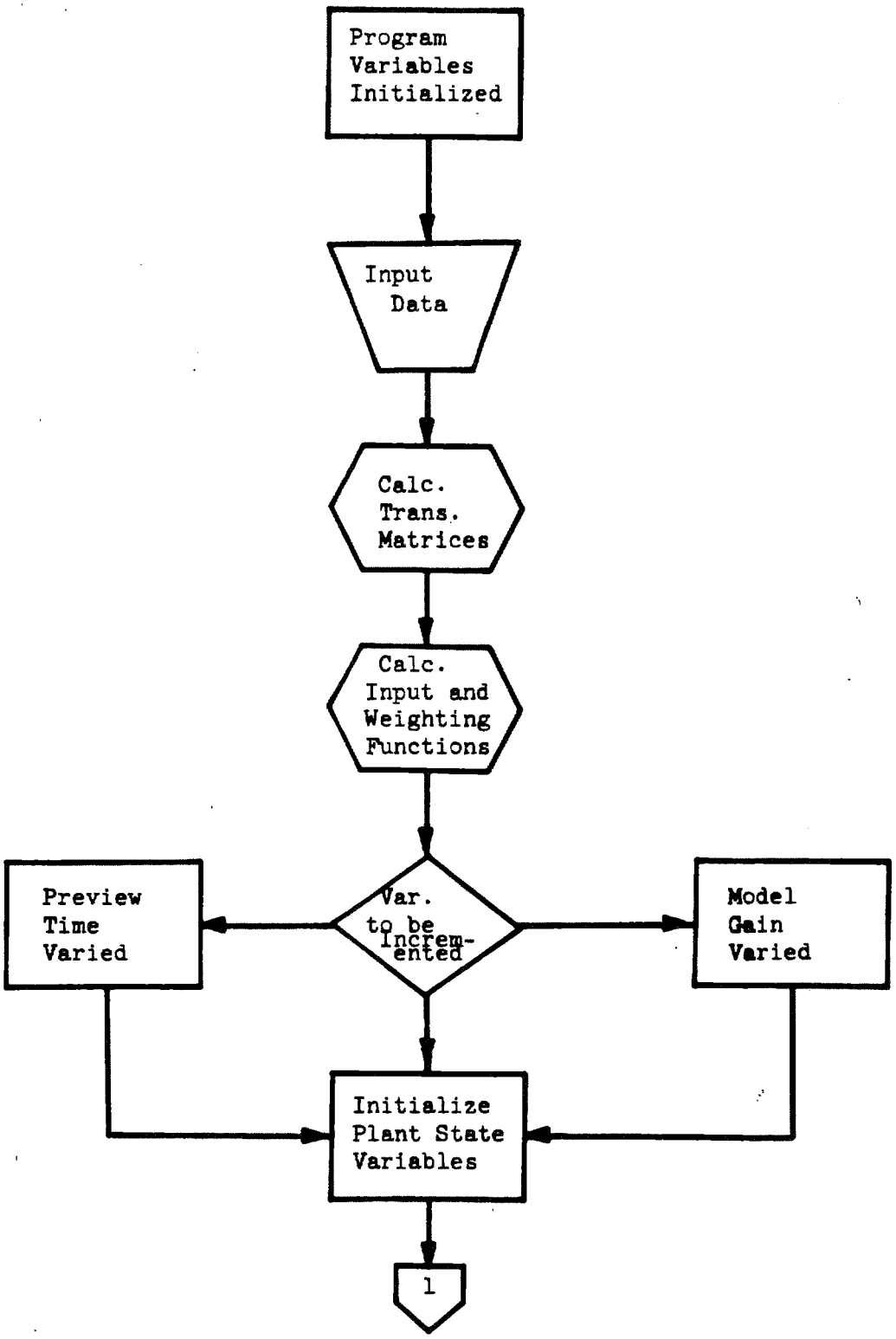
$$= 15*(20 + 20!) \text{ calculations must be made.}$$

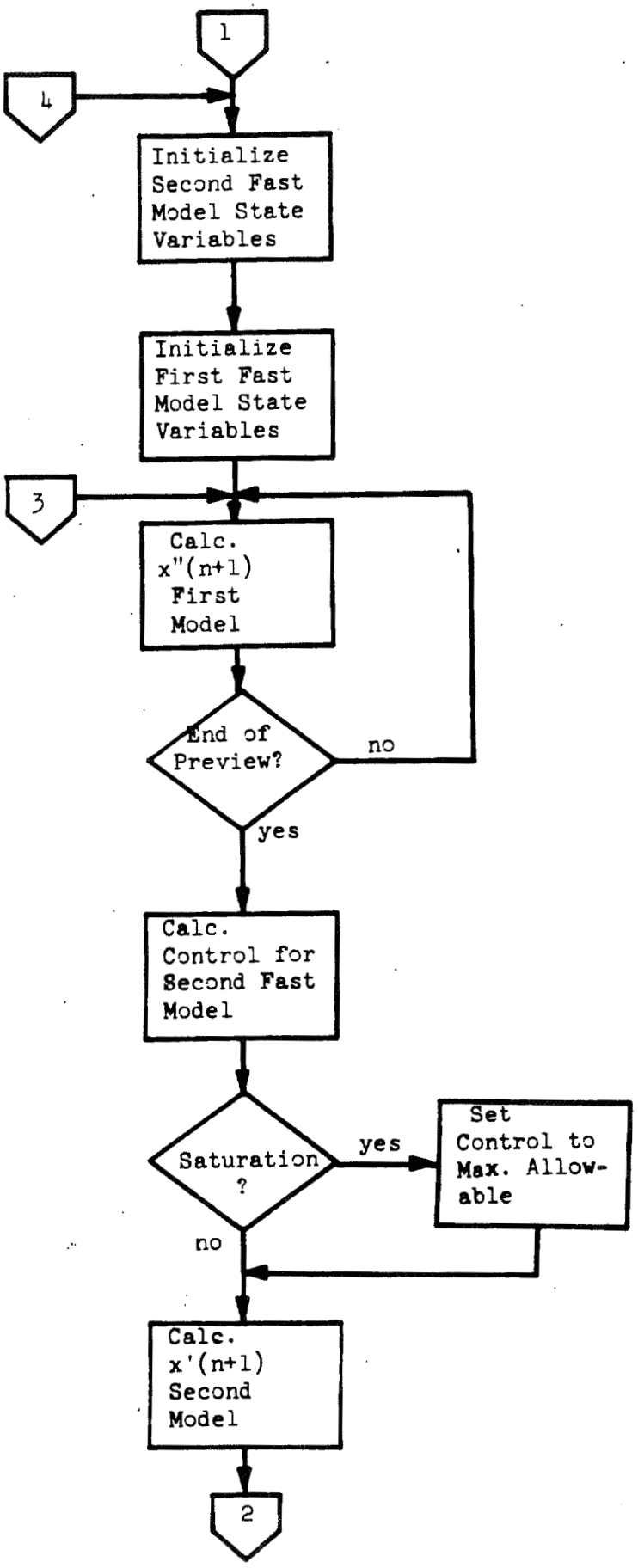
Notice, that if only one fast model is used, the factorial is eliminated.

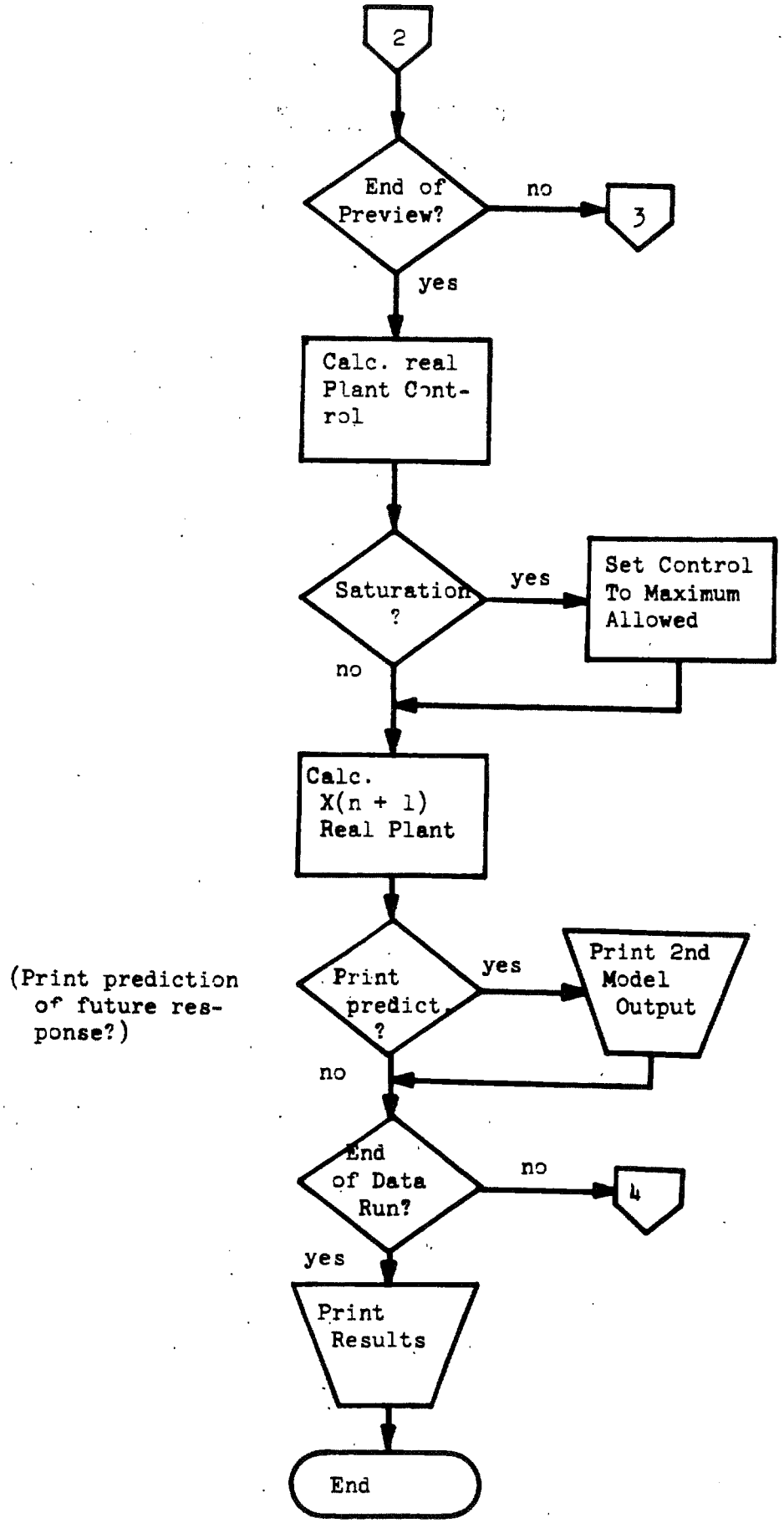
This would require a serious look at the economic gains to be received from the second fast model if a program were to be written for a practical problem. In general, using the I.B.M. 7094 computer the existing programs will run in real time if preview does not exceed 3 units.

All programming was done in the M.A.D. language using the Compatible Time Sharing System of the M.I.T. Computation Center. Data was taken using batch processing at the same center.

Figure B-1. FLOWCHART OF SIMULATION PROGRAM FOR PREVIEW CONTROL MODEL USING TWO FAST MODELS




1
4
Initialize Second Fast Model State Variables
Initialize First Fast Model State Variables
3
Calc. x"(n+1) First Model
End of Preview?
no
yes
Calc. Control for Second Fast Model
Saturation ?
yes
Set Control to Max. Allow-able
no
Calc. x'(n+1) Second Model
2


2
End of Preview?
no
3
yes
Calc. real Plant Control
Saturation ?
yes
Set Control To Maximum Allowed
no
Calc. X(n + 1) Real Plant
(Print prediction of future response?)
Print predict. ?
yes
Print 2nd Model Output
no
End of Data Run?
no
4
yes
Print Results
End

APPENDIX C DATA

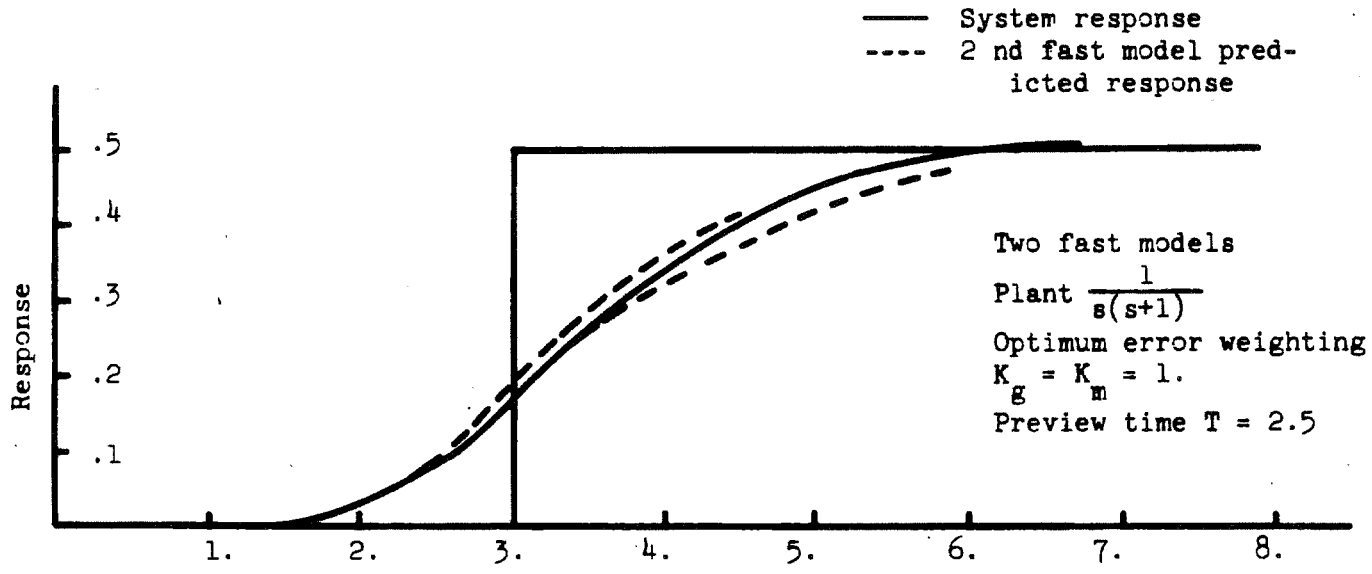


Figure C-1

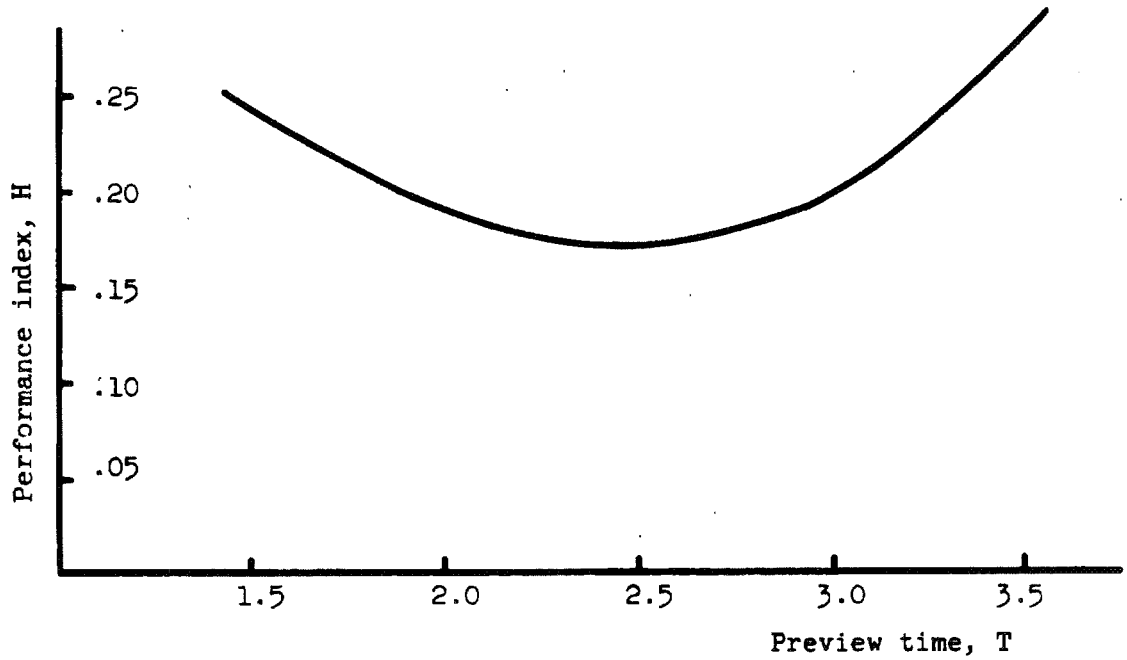


Figure C-2

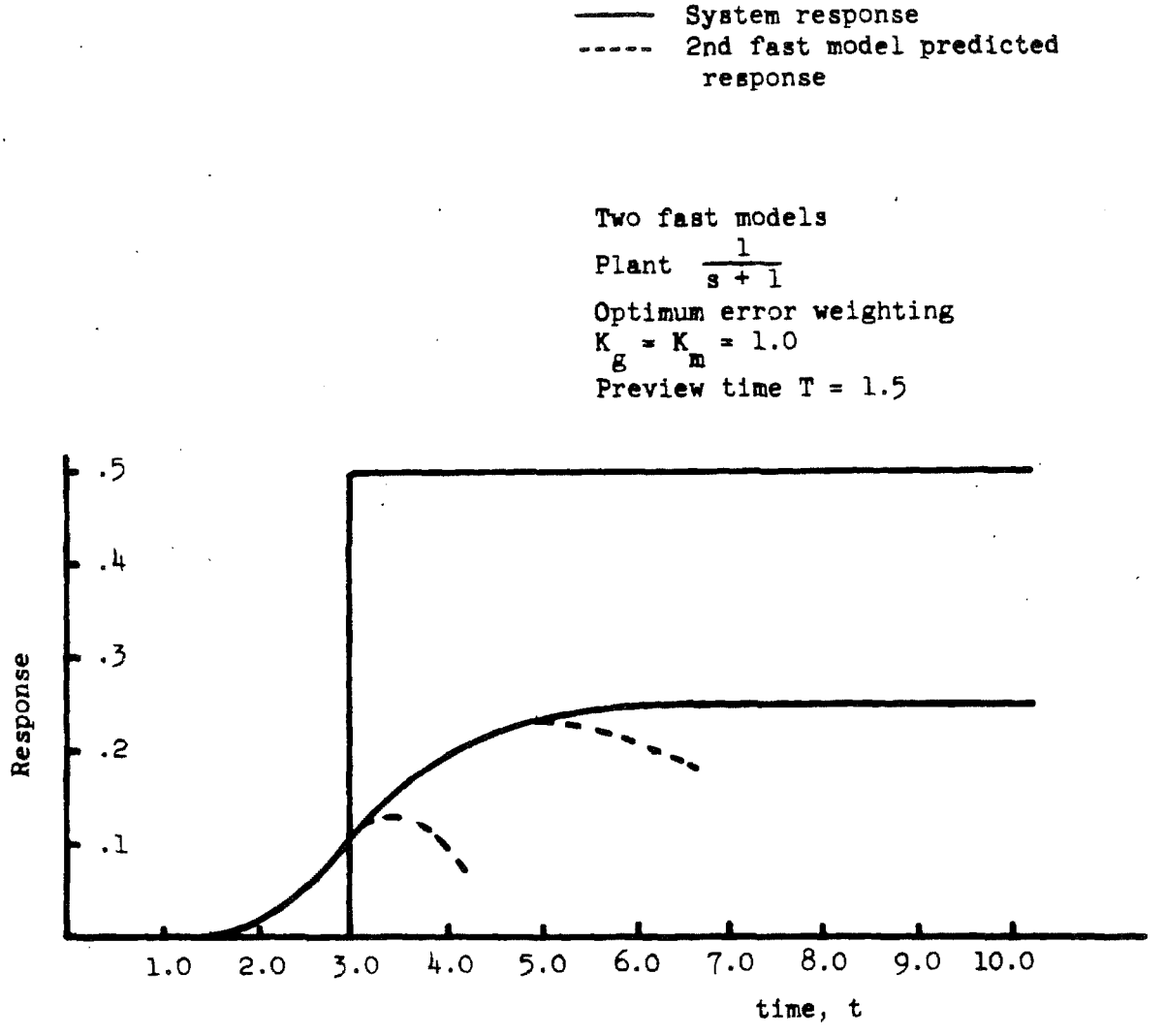

System response
2nd fast model predicted response
Two fast models
Plant $\frac{1}{s+1}$
Optimum error weighting
$K_g = K_m = 1.0$
Preview time T = 1.5
Response
.5
.4
.3
.2
.1
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0
10.0
time, t

Figure C-3

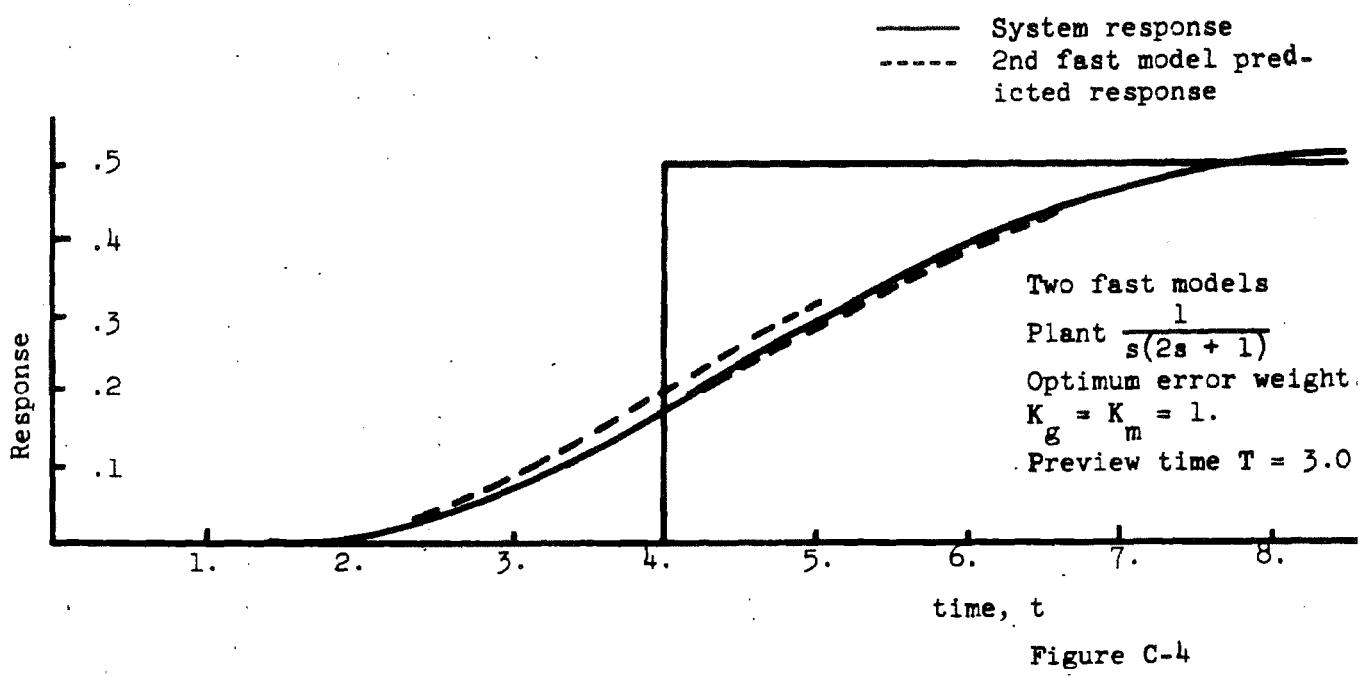


Figure C-4

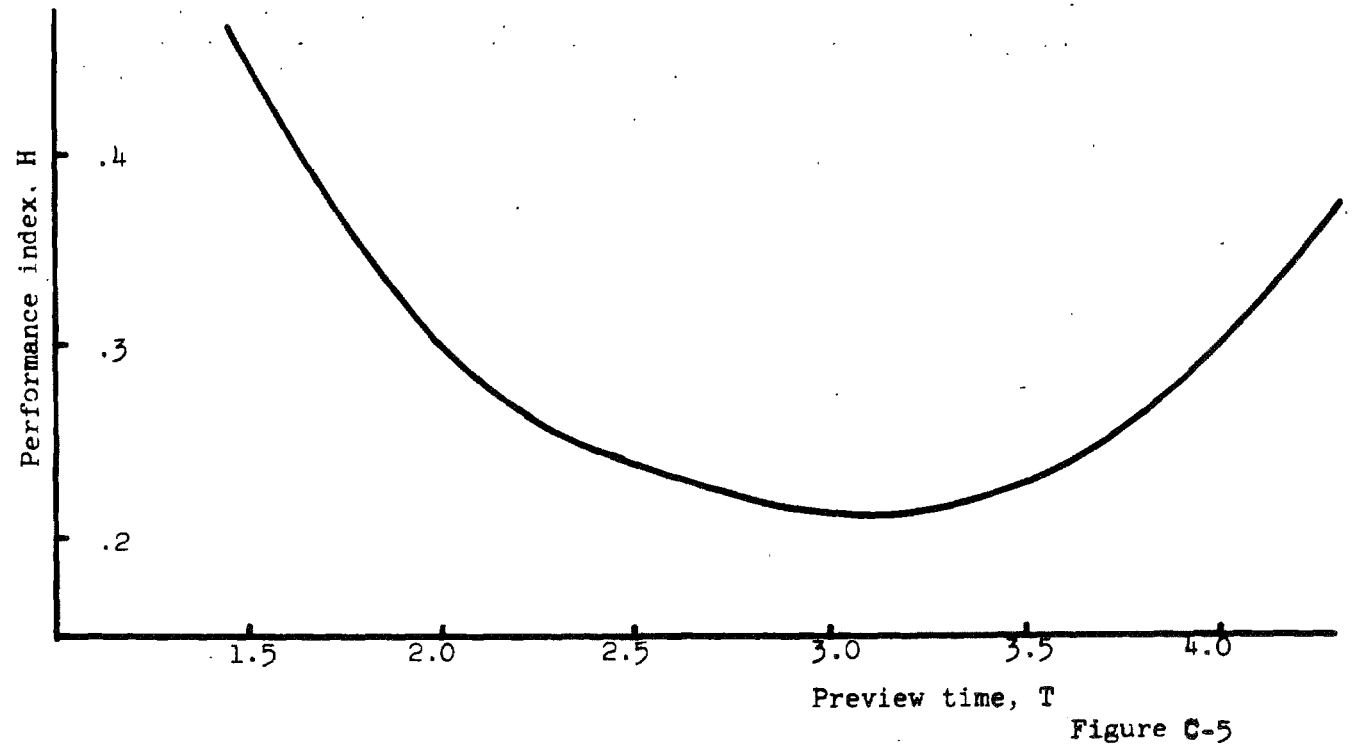


Figure C-5

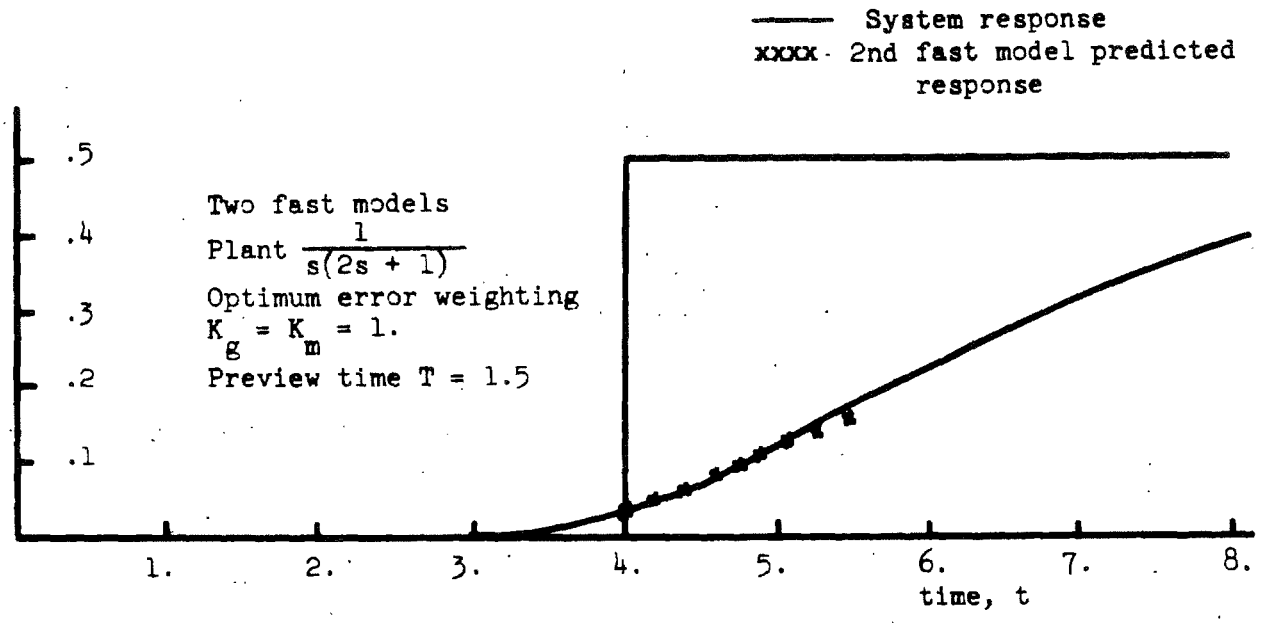


Figure C-6

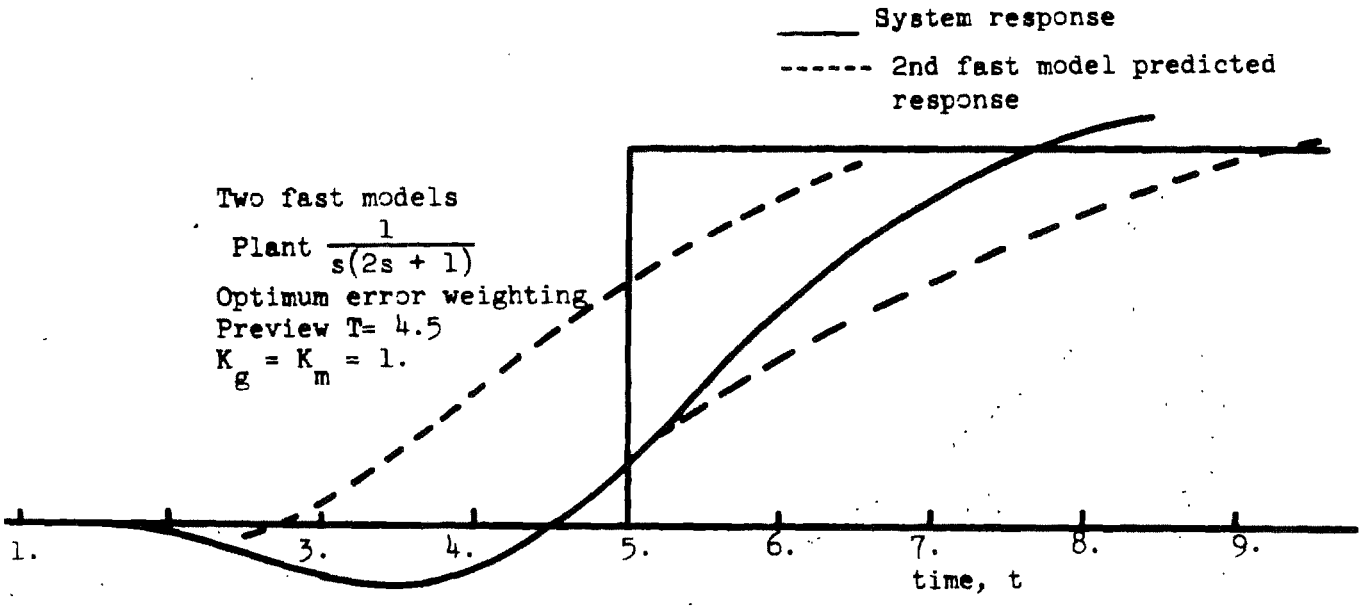


Figure C-7

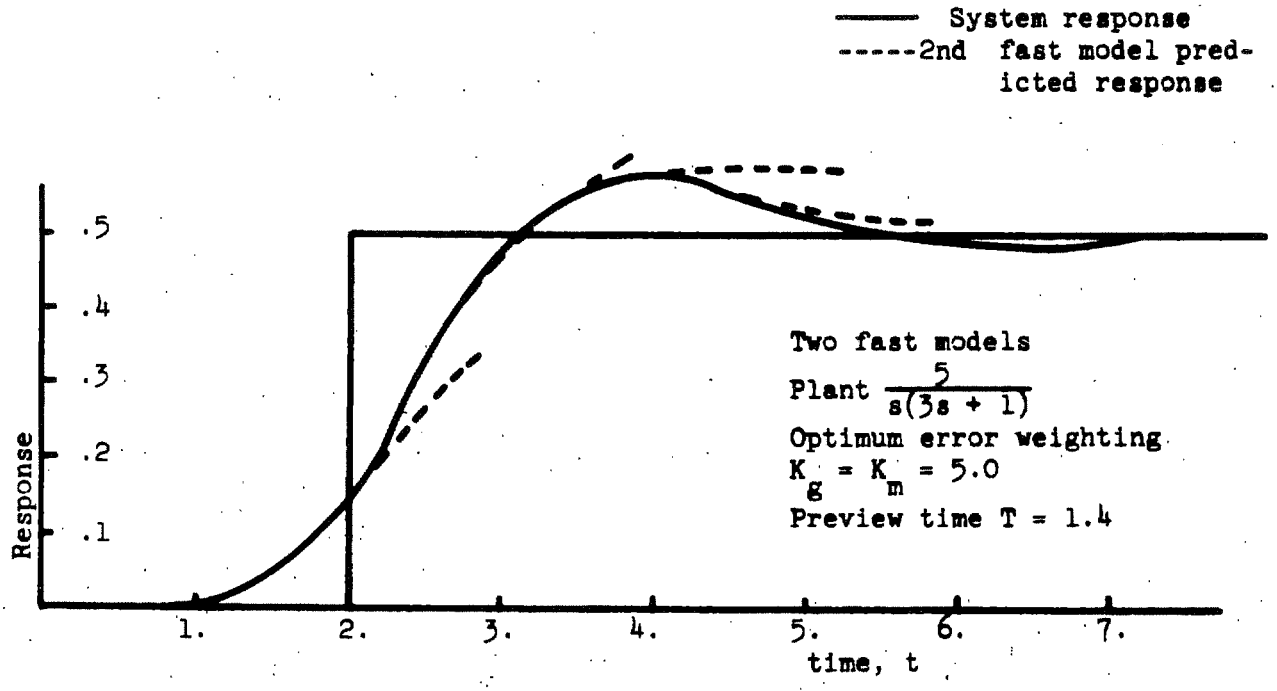


Figure C-8

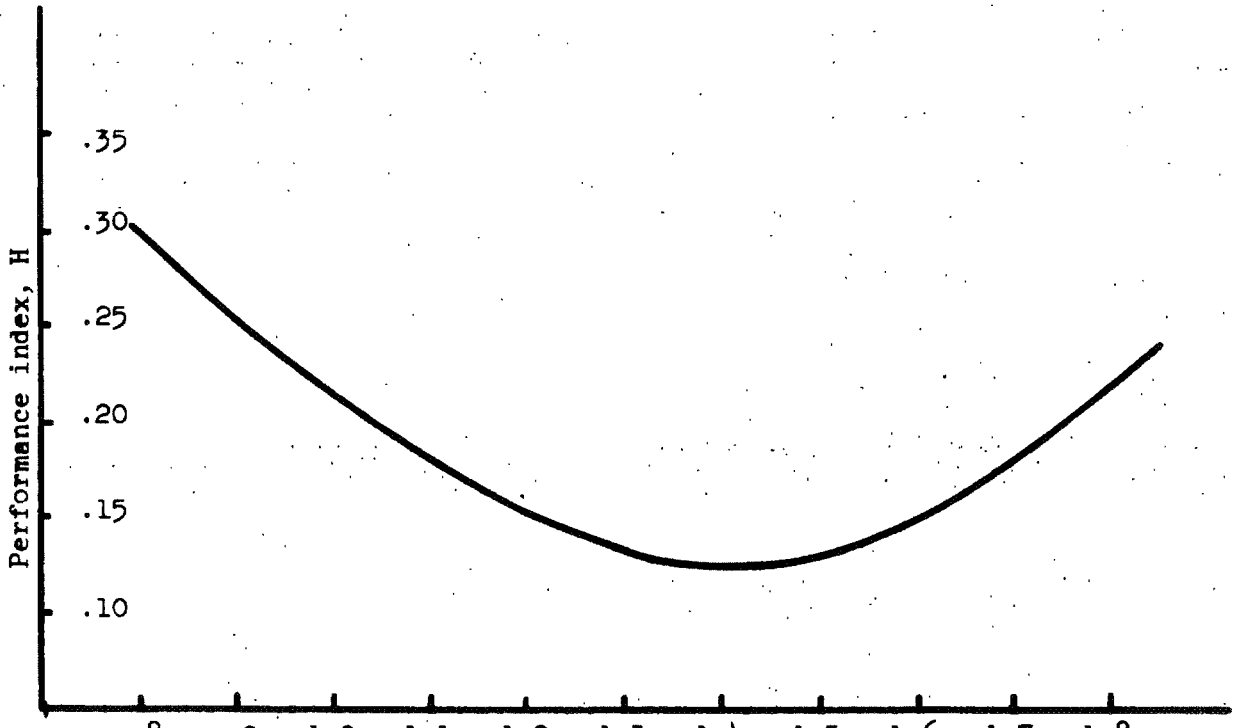


Figure C-9

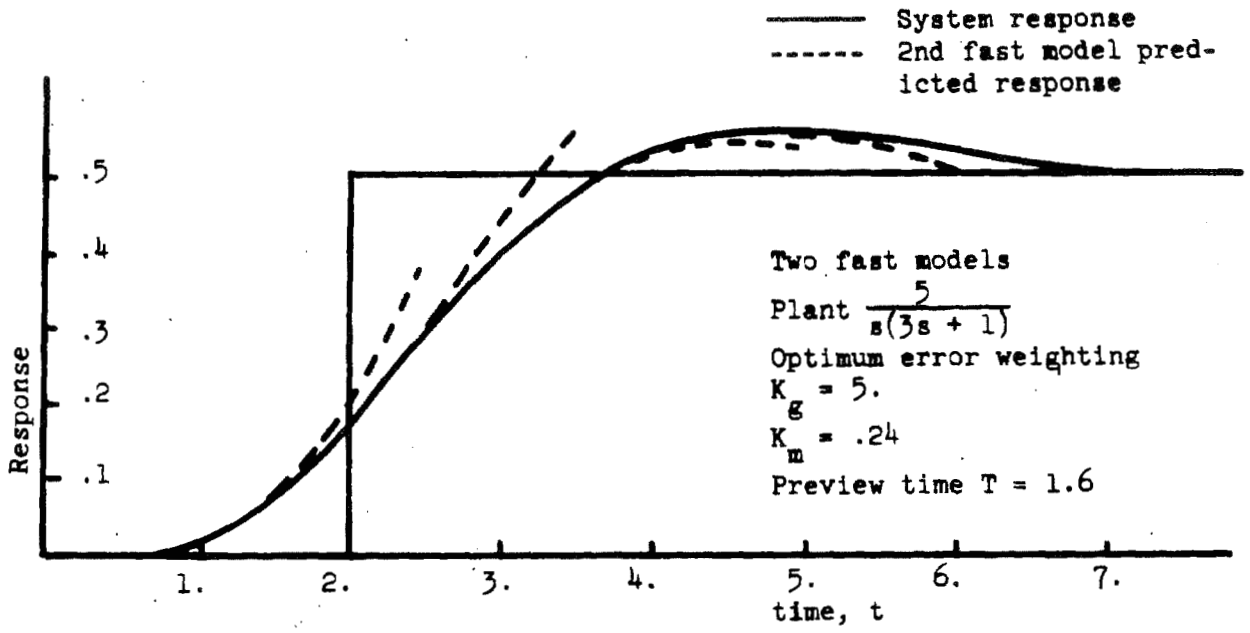


Figure C-10

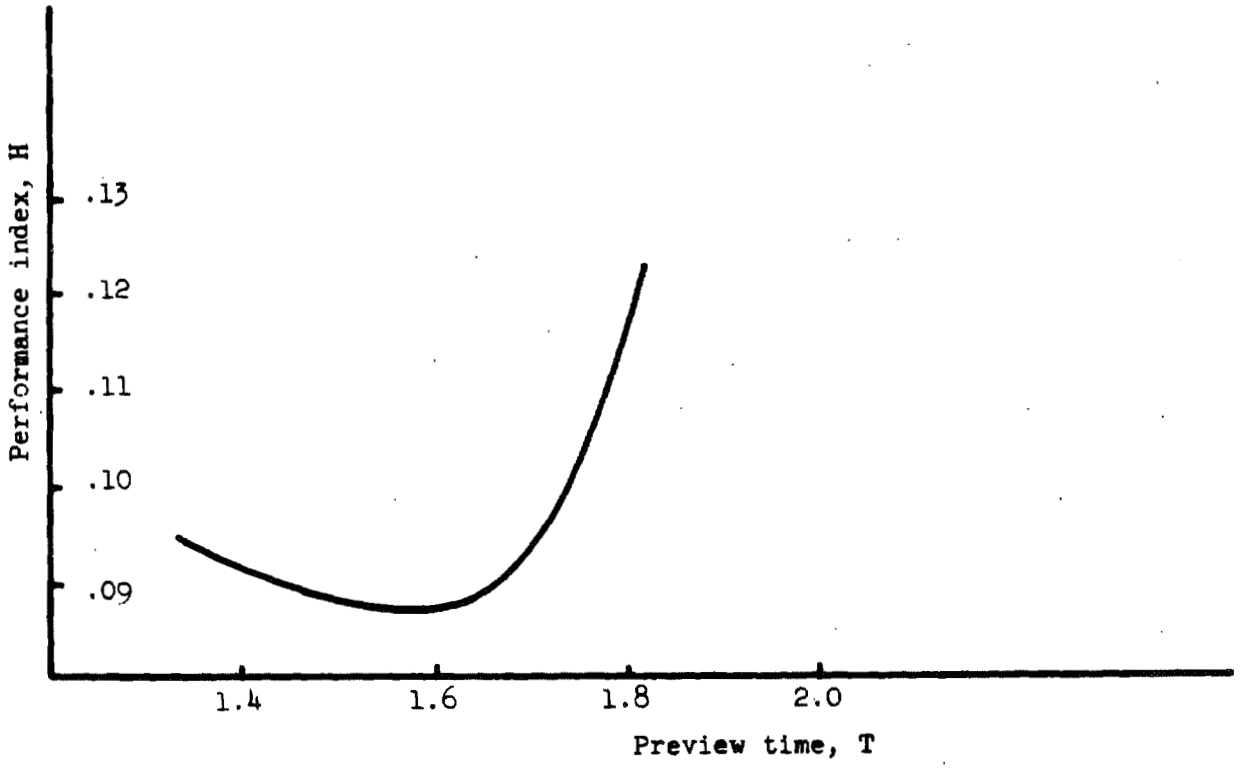
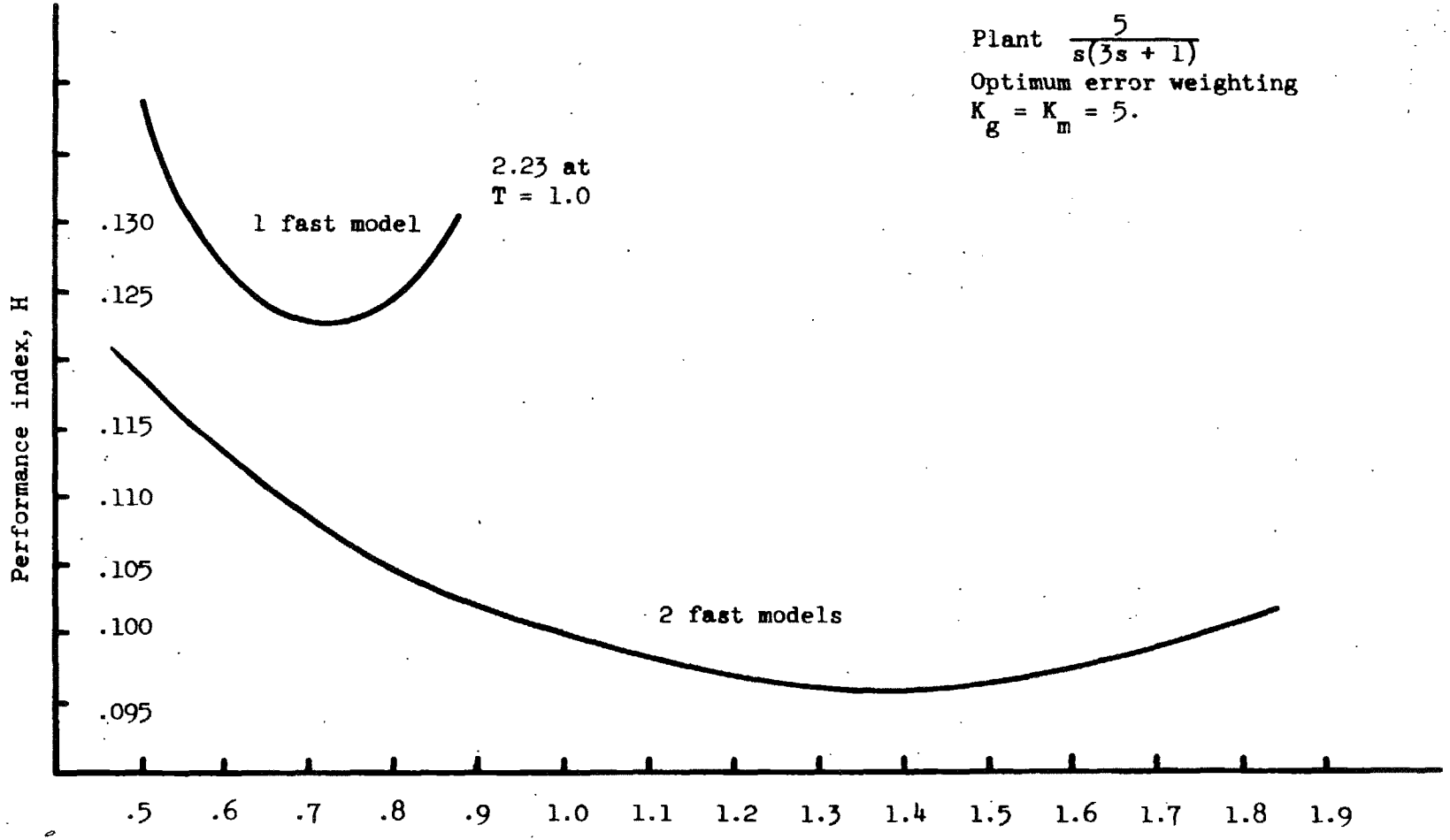


Figure C-11



Preview vs. H, Comparison one and two fast models

Figure C-12

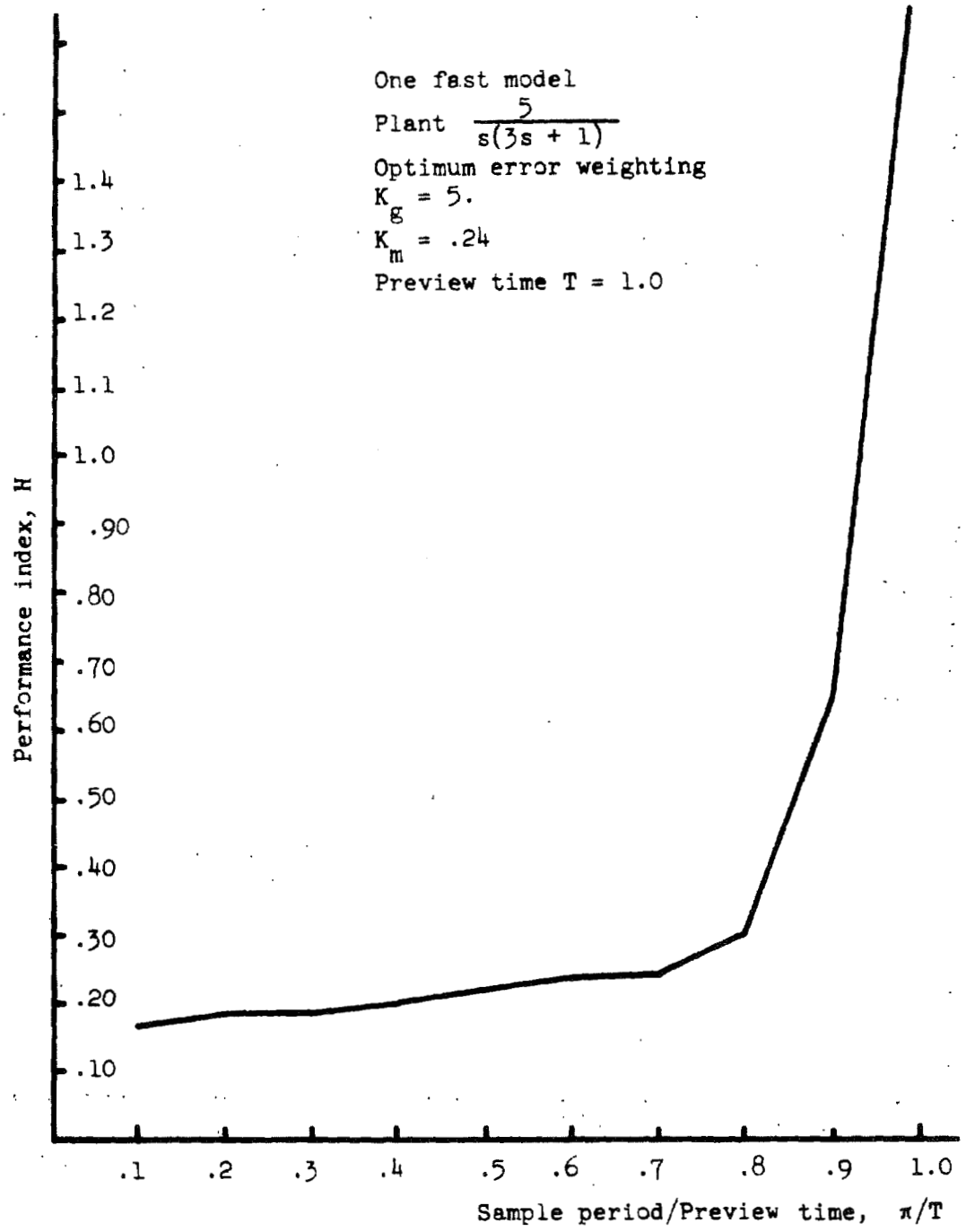

One fast model
Plant $\frac{5}{s(3s + 1)}$
Optimum error weighting
$K_g = 5.$
$K_m = .24$
Preview time T = 1.0
Performance index, H
1.4
1.3
1.2
1.1
1.0
.90
.80
.70
.60
.50
.40
.30
.20
.10
.1
.2
.3
.4
.5
.6
.7
.8
.9
1.0
Sample period/Preview time, π/T

Figure C-13

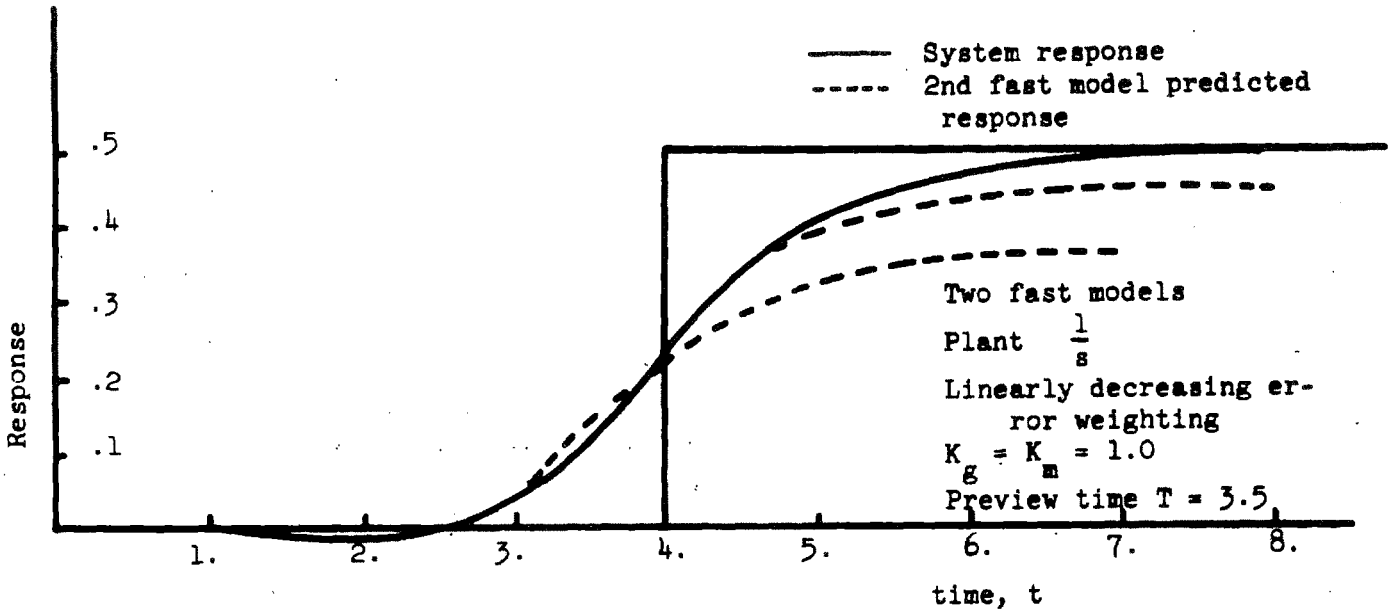


Figure C-14

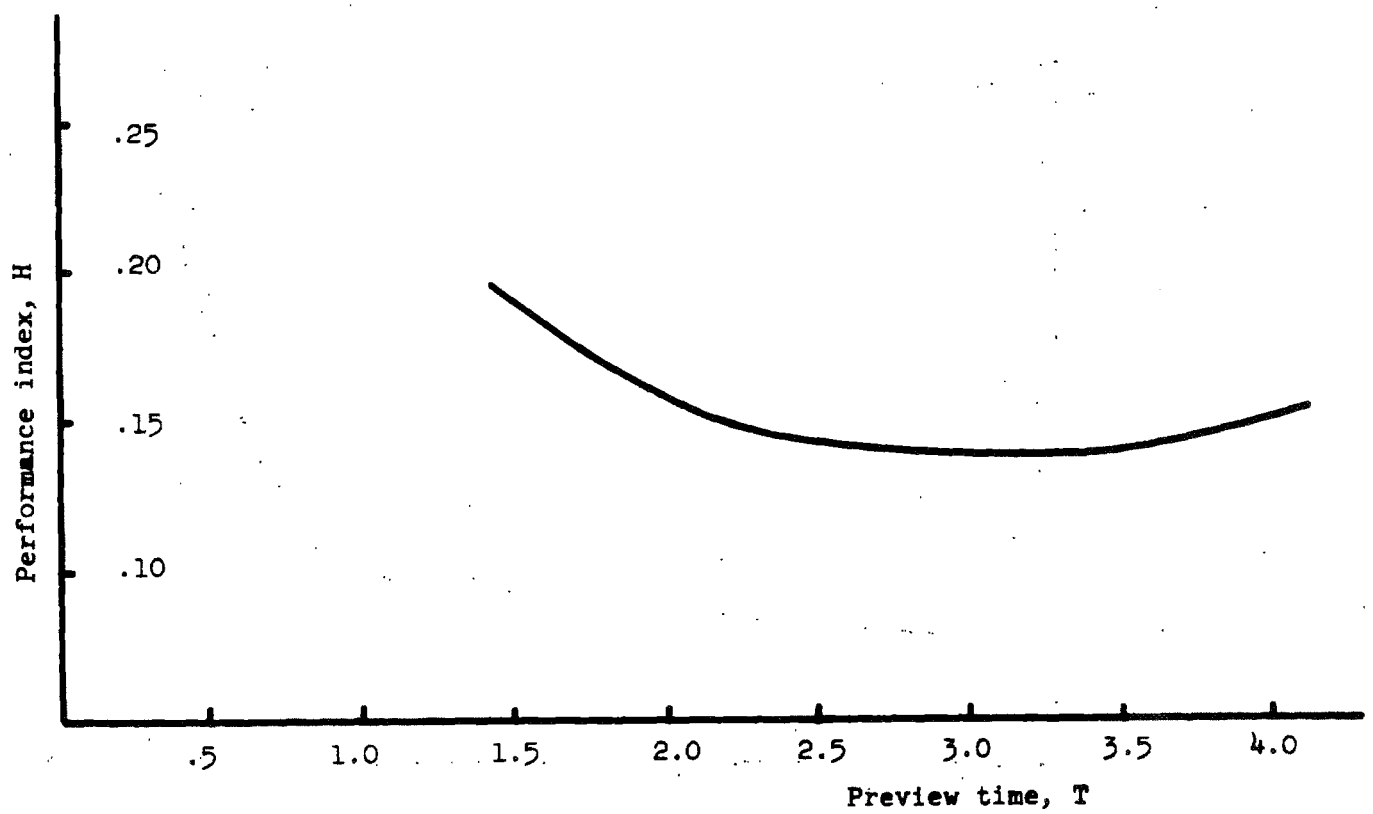


Figure C-15

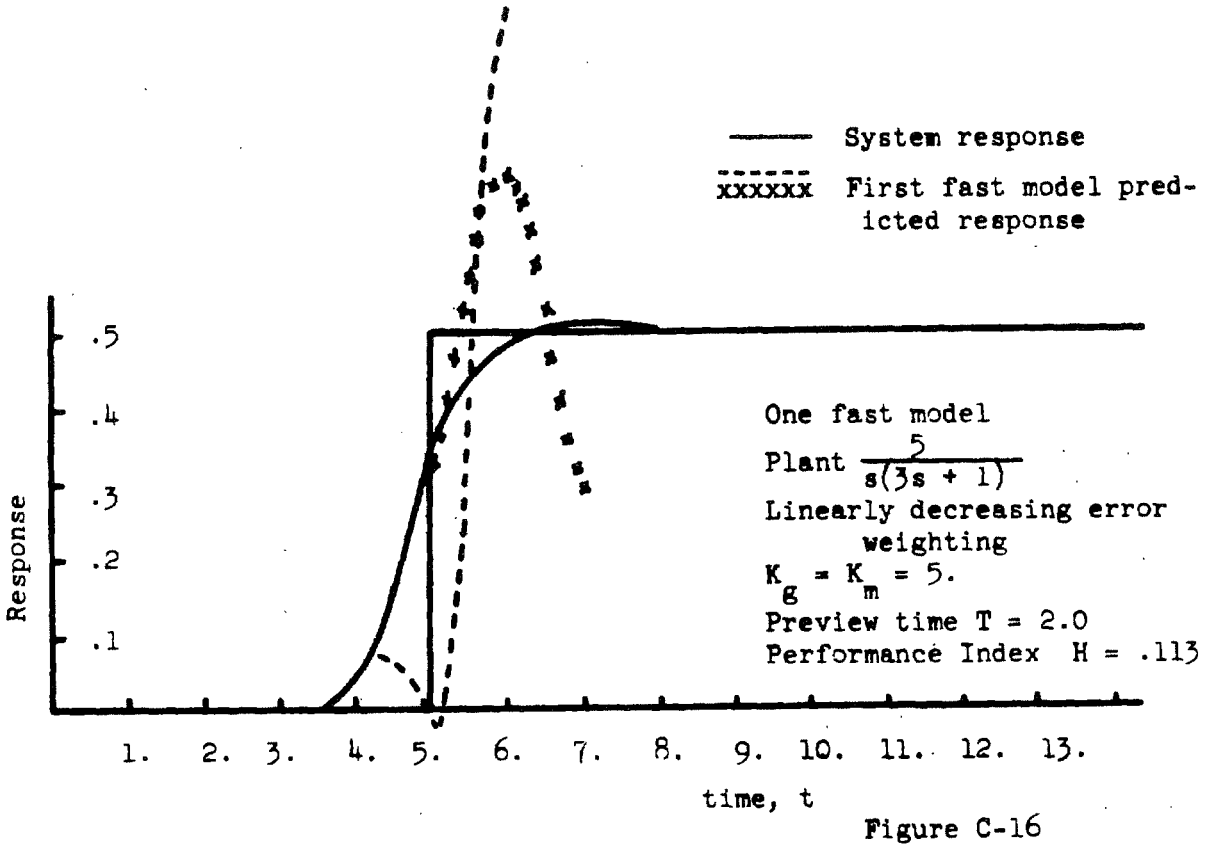


Figure C-16

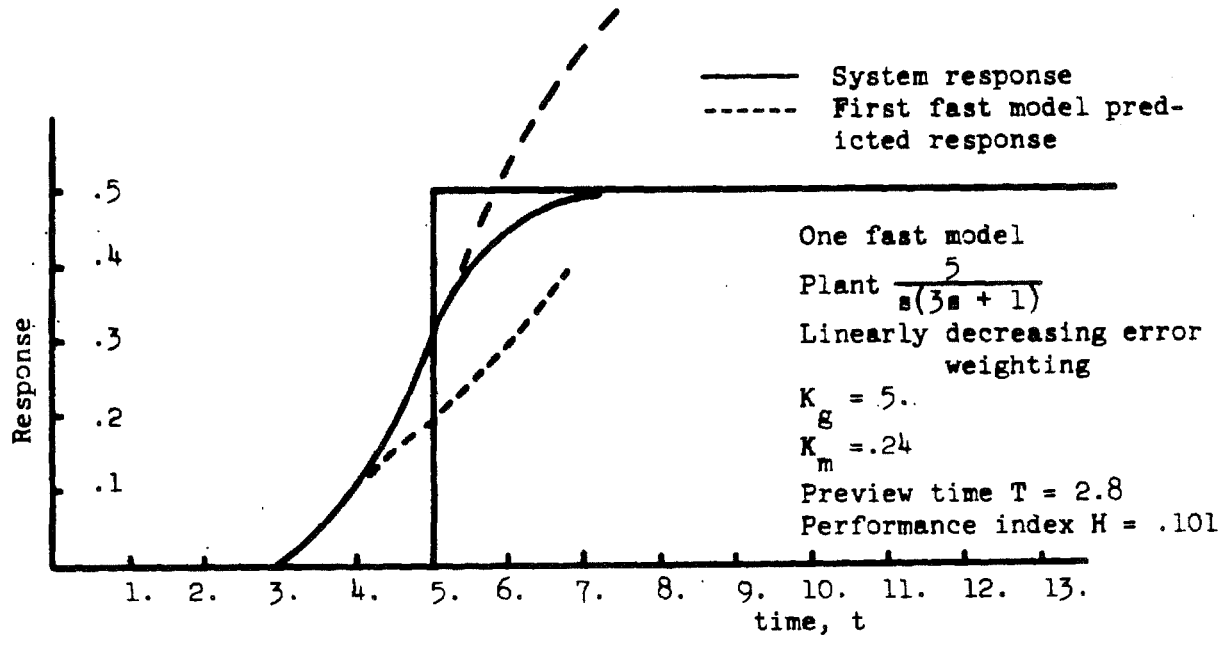


Figure C-17

Plant $\frac{5}{s(3s + 1)}$

$K_g = 5.$

$K_m = .24$

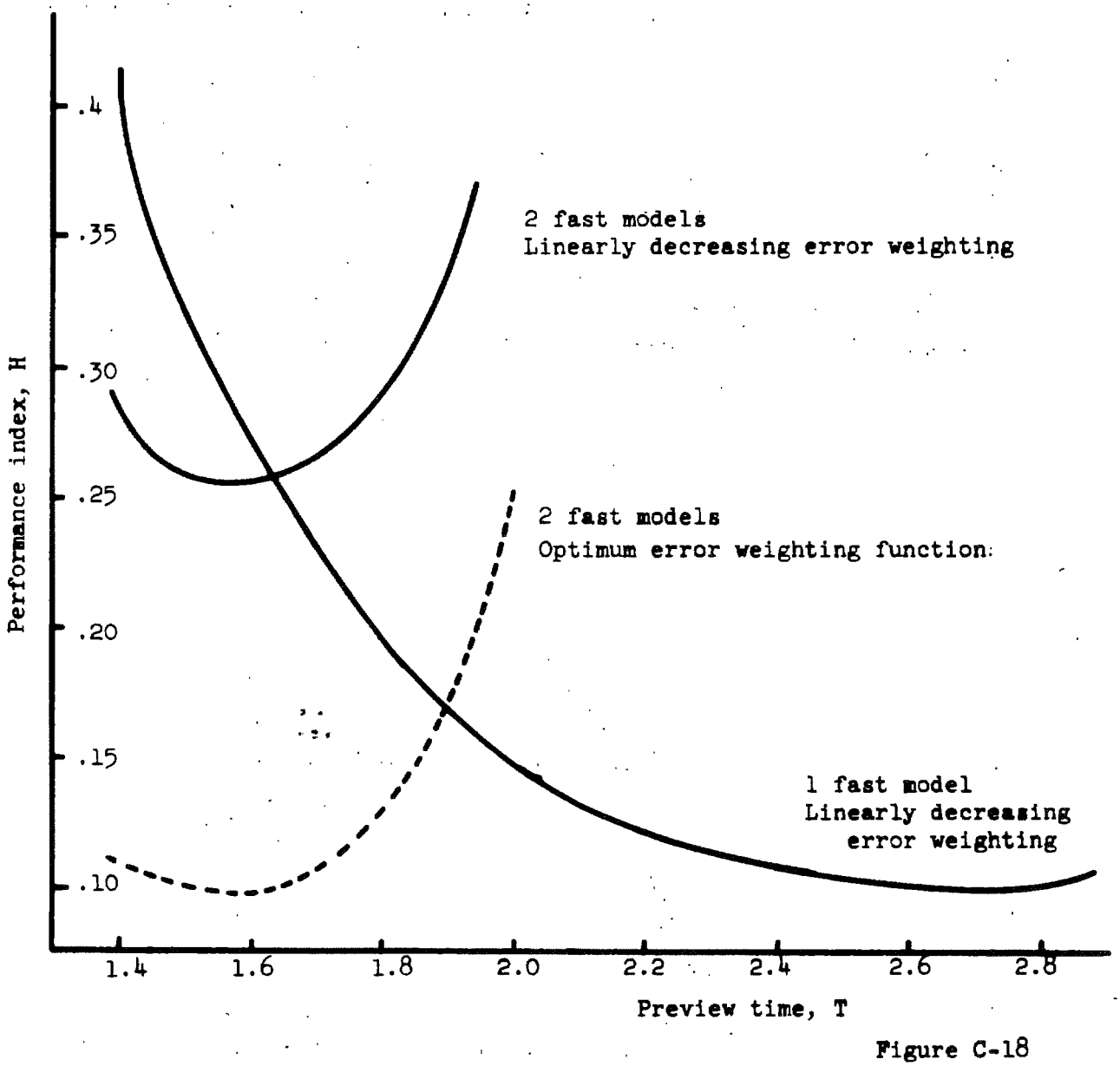


Figure C-18

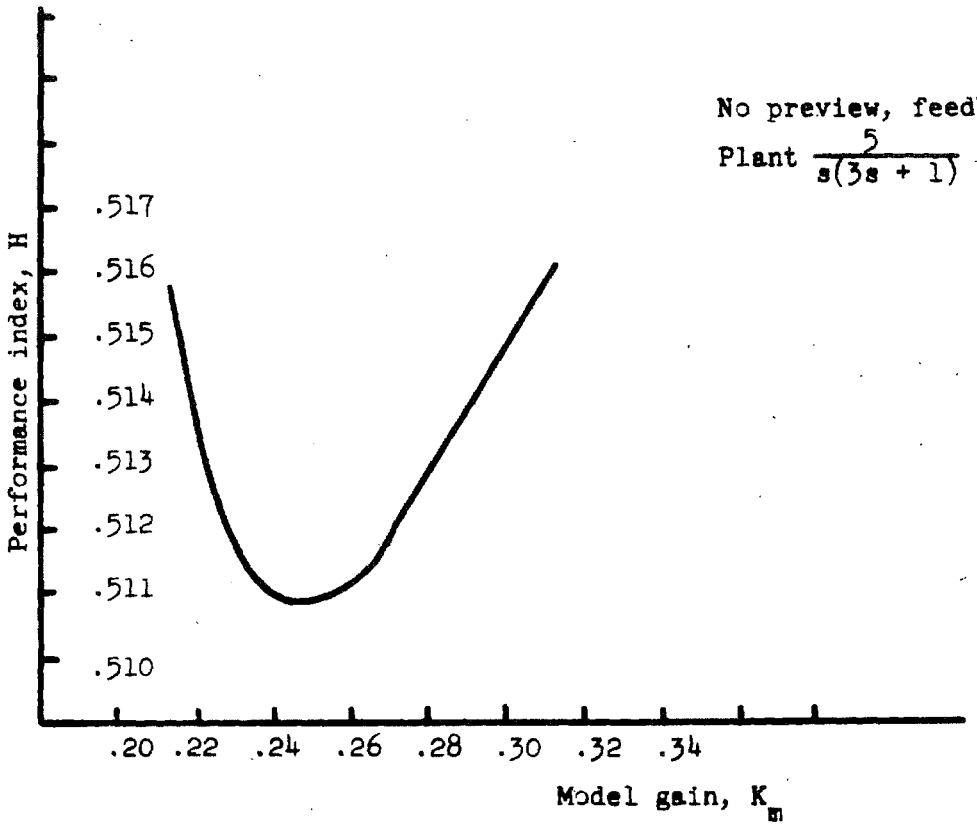
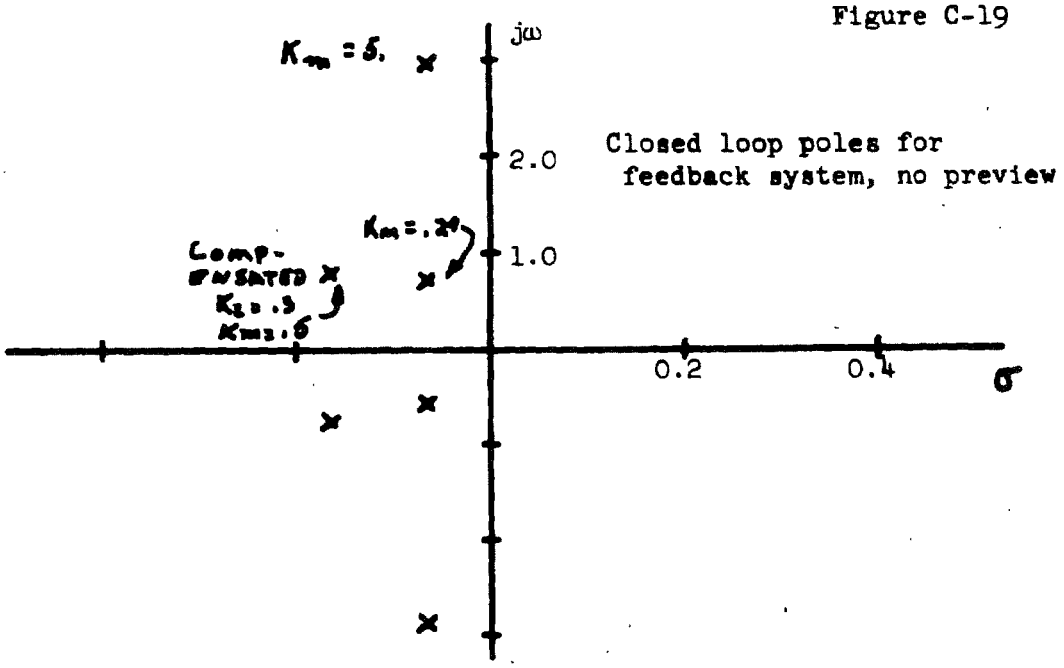


Figure C-19



Closed loop poles for feedback system, no preview

Figure C-20

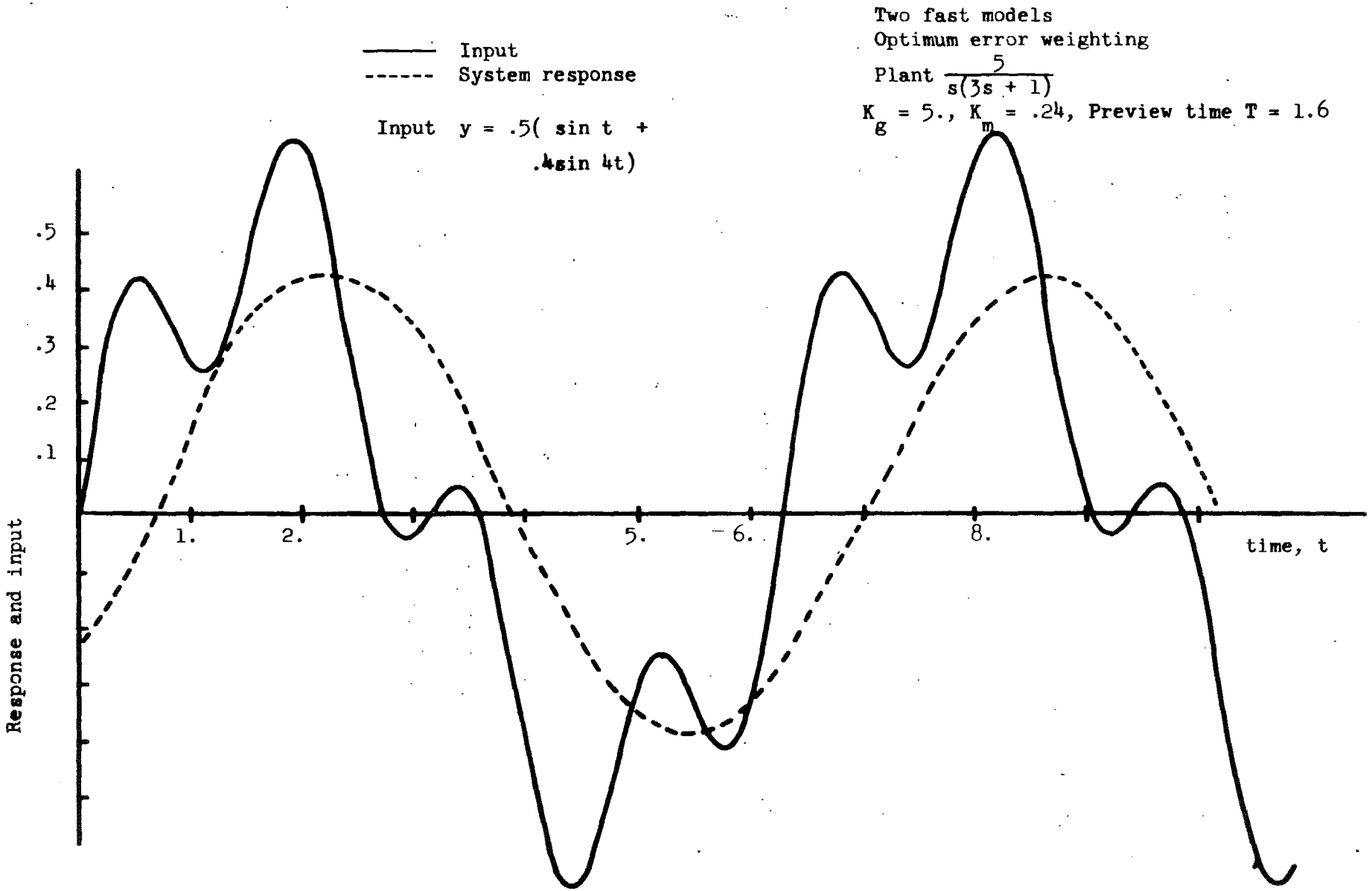

Input
System response
Input y = .5(sin t + .4sin 4t)
Two fast models
Optimum error weighting
Plant 5/s(3s + 1)
K_g = 5., K_m = .24, Preview time T = 1.6
.5
.4
.3
.2
.1
1.
2.
5.
6.
8.
time, t
Response and input

Figure C-21

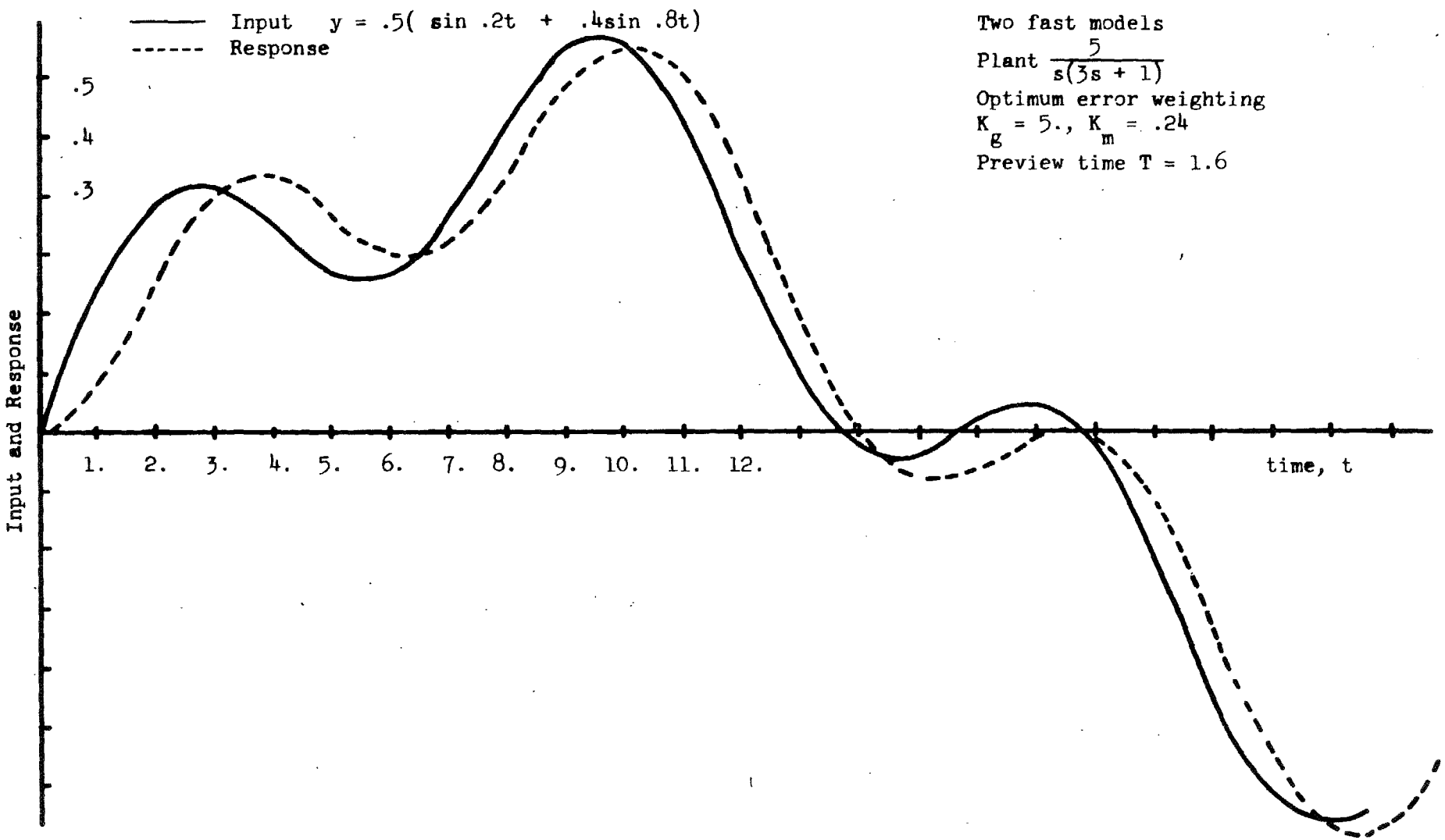
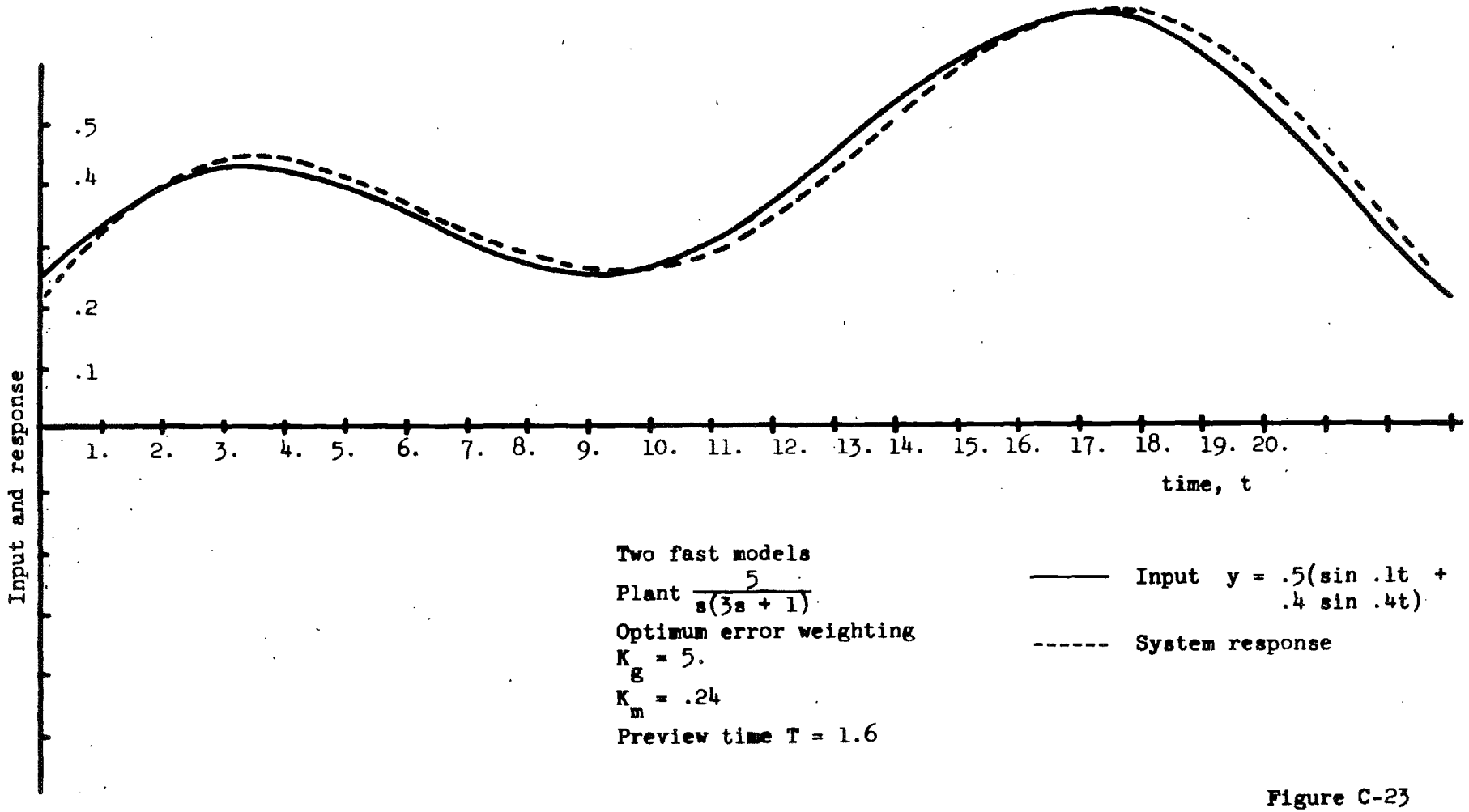

Input y = .5(sin .2t + .4sin .8t)
Response
Two fast models
Plant $\frac{5}{s(3s + 1)}$
Optimum error weighting
$K_g = 5.$, $K_m = .24$
Preview time T = 1.6
.5
.4
.3
Input and Response
1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12.
time, t

Figure C-22



Two fast models

Plant $\frac{5}{s(3s + 1)}$

Optimum error weighting

$K_g = 5.$

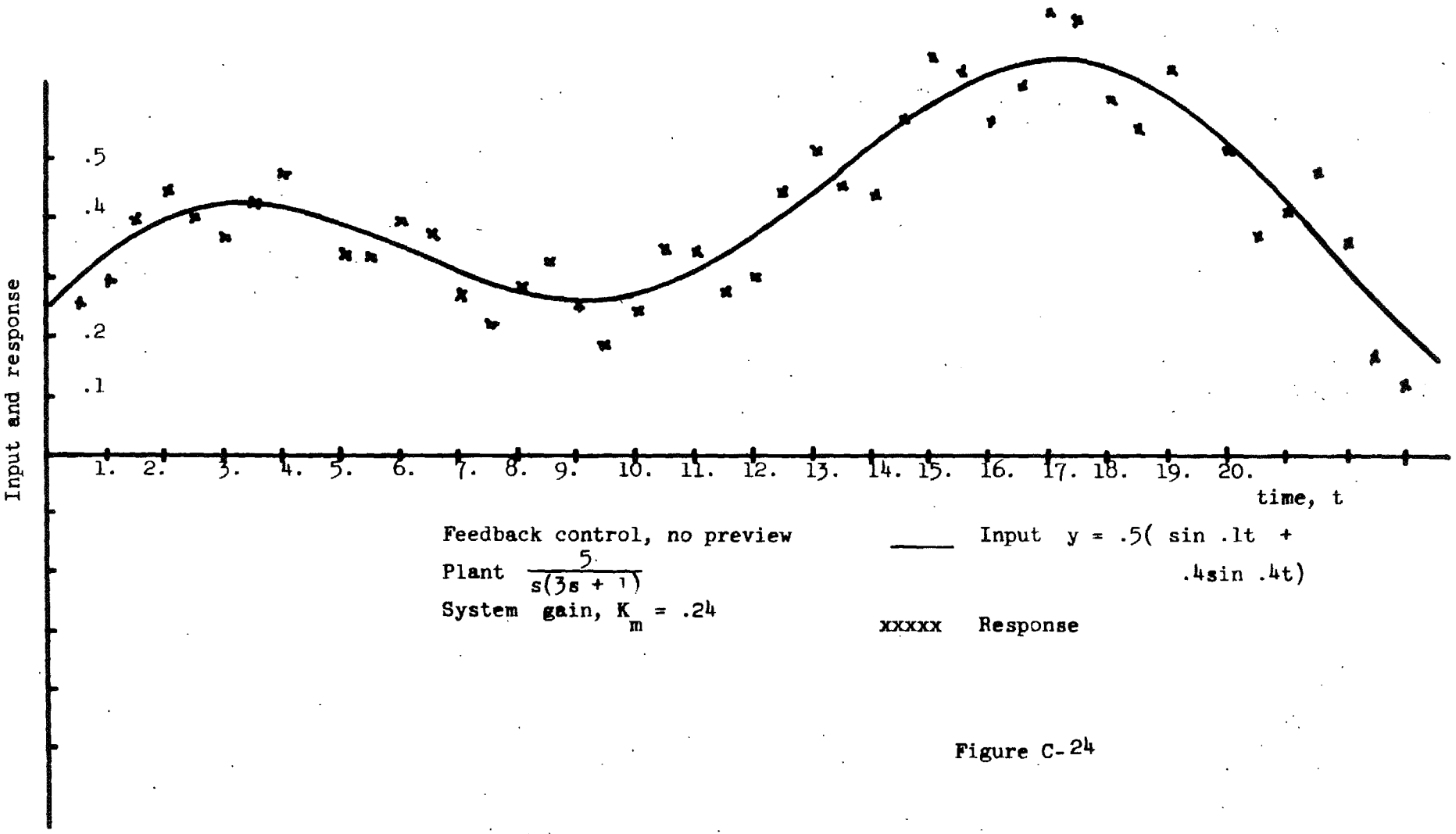
$K_m = .24$

Preview time T = 1.6

——— Input $y = .5(\sin .1t + .4 \sin .4t)$

------ System response

Figure C-23



Feedback control, no preview

Plant $\frac{5}{s(3s + 1)}$

System gain, $K_m = .24$

____ Input $y = .5(\sin .1t + .4\sin .4t)$

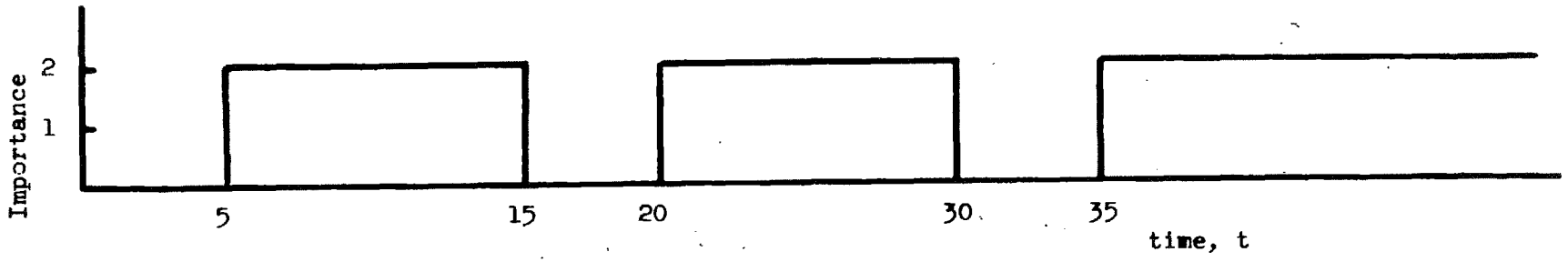
xxxxx Response

Figure C-24

Two fast models $K_g = 5.$ Preview time T = 1.6

Plant $\frac{5}{s(3s + 1)}$ $K_m = .24$

Optimum weighting function (optimum for uniformly weighted error)



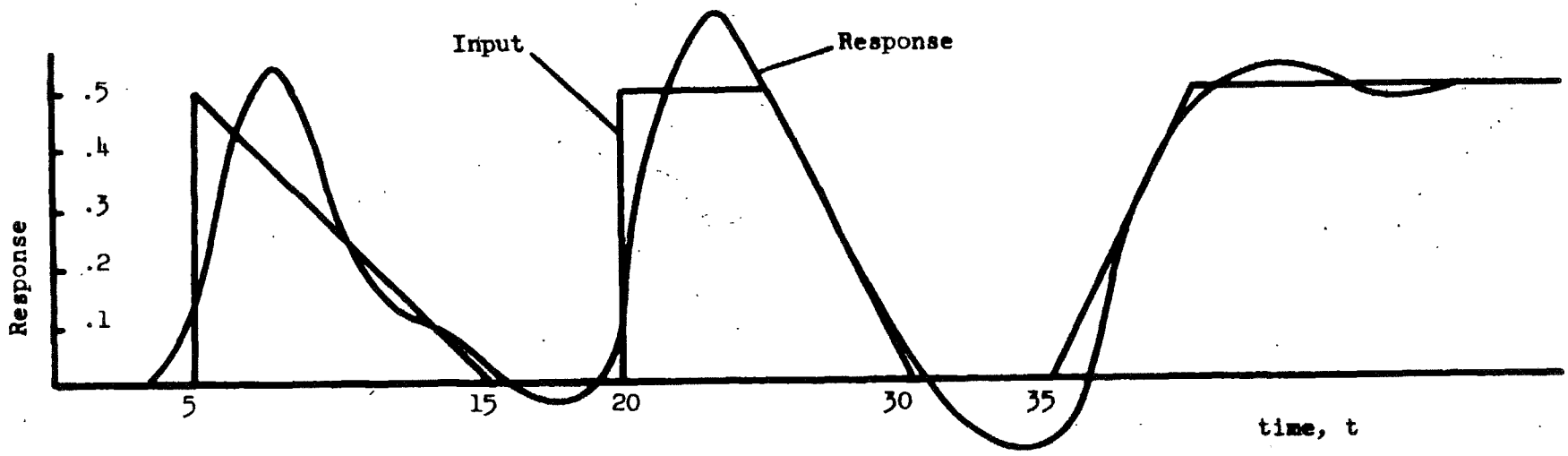


Figure C-25

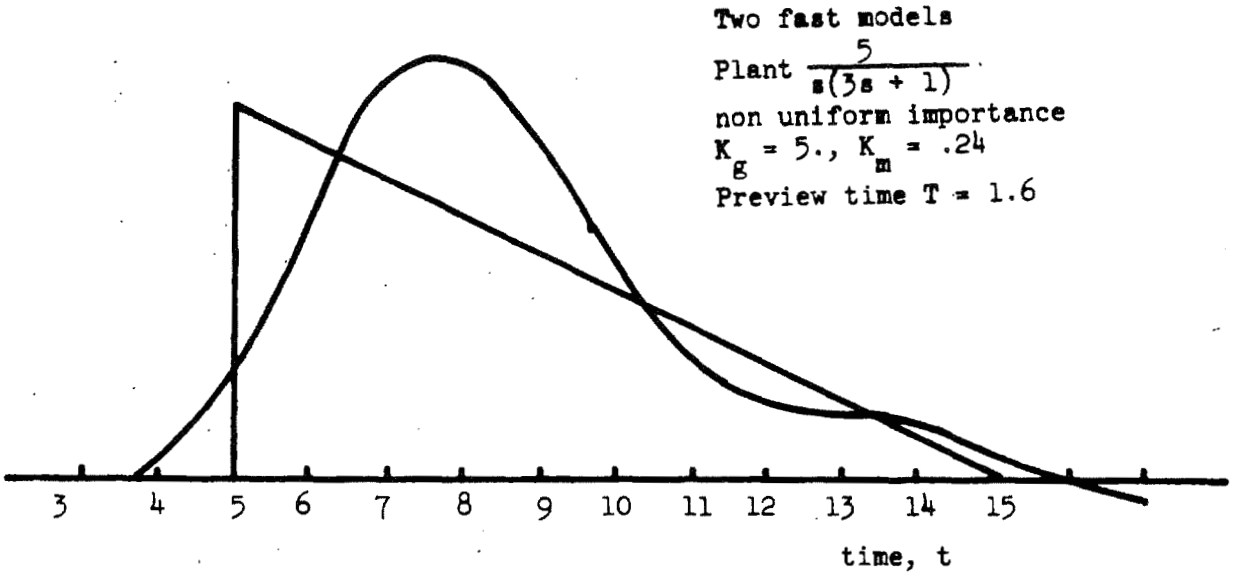


Figure C-25

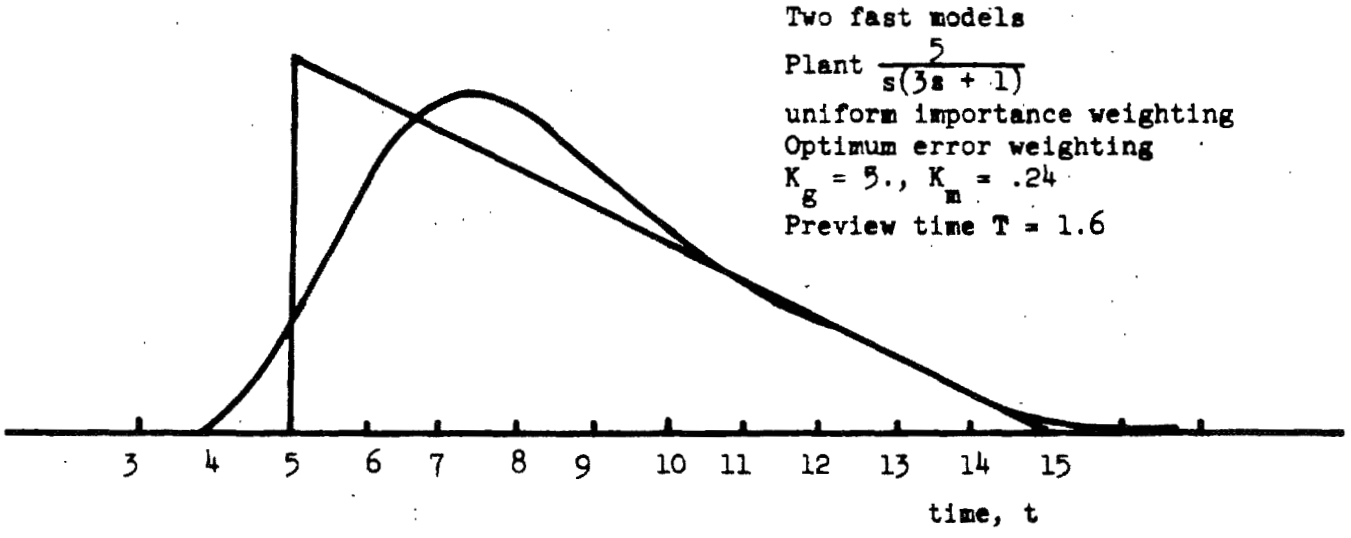

Two fast models
Plant $\frac{5}{s(3s+1)}$
uniform importance weighting
Optimum error weighting
$K_g = 5.$, $K_m = .24$
Preview time T = 1.6
3 4 5 6 7 8 9 10 11 12 13 14 15
time, t

Figure 26

Two fast models

Plant $\frac{1}{s}$

Flat error weighting

Non-uniform importance weighting

$K_g = K_m = 1.$

Preview time T = 1.5

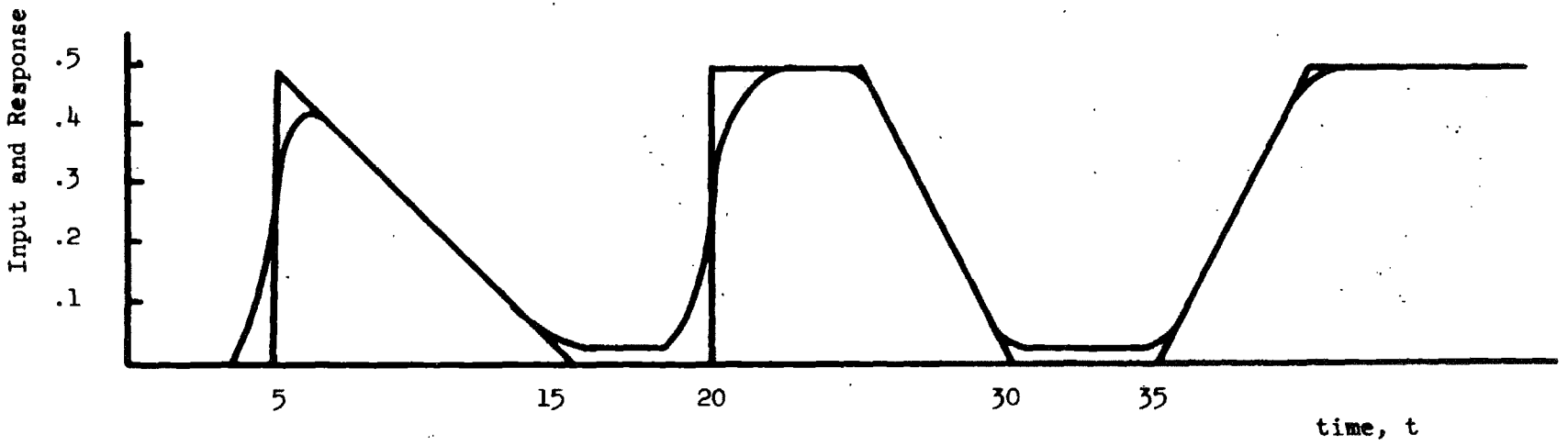
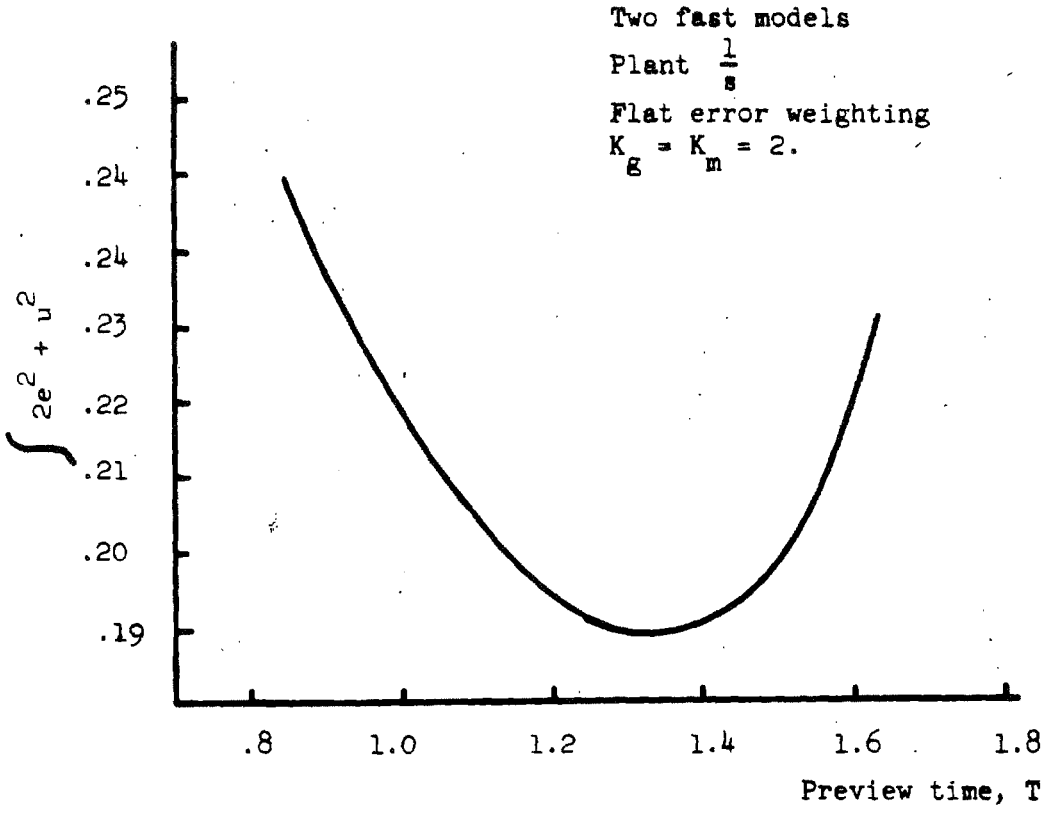


Figure C-27


Two fast models
Plant $\frac{1}{s}$
Flat error weighting
$K_g = K_m = 2.$
$\int 2e^2 + u^2$
.25
.24
.24
.23
.22
.21
.20
.19
.8
1.0
1.2
1.4
1.6
1.8
Preview time, T

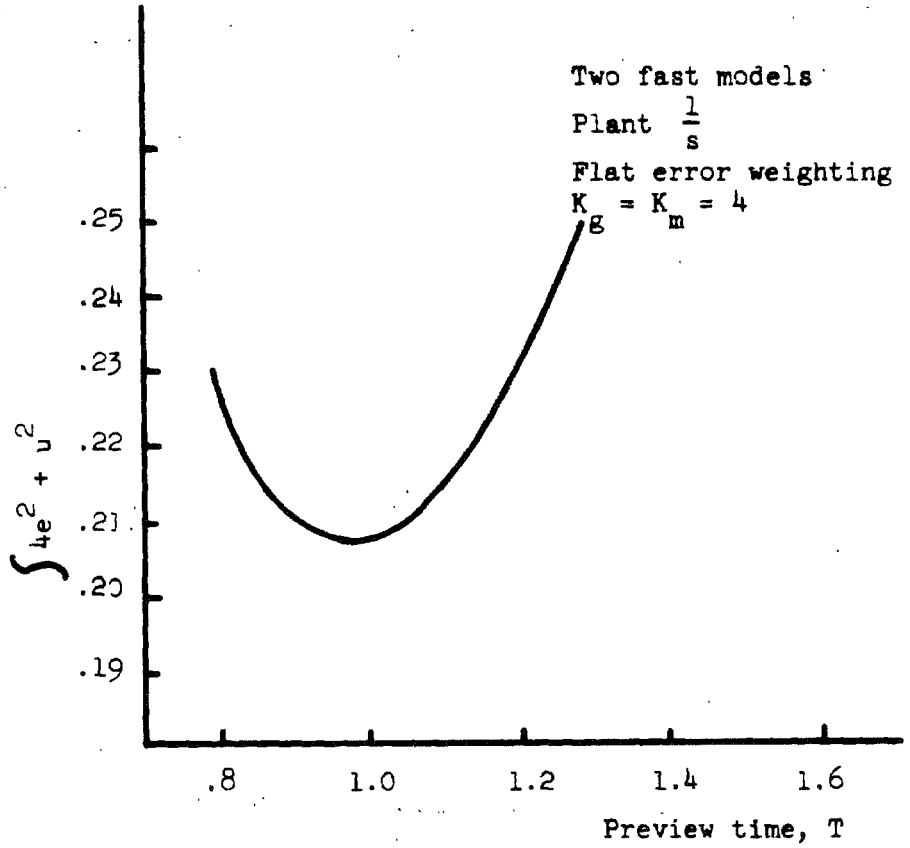

Two fast models
Plant $\frac{1}{s}$
Flat error weighting
$K_g = K_m = 4$
$\int 4e^2 + u^2$
.25
.24
.23
.22
.21
.20
.19
.8
1.0
1.2
1.4
1.6
Preview time, T

Figure C-28

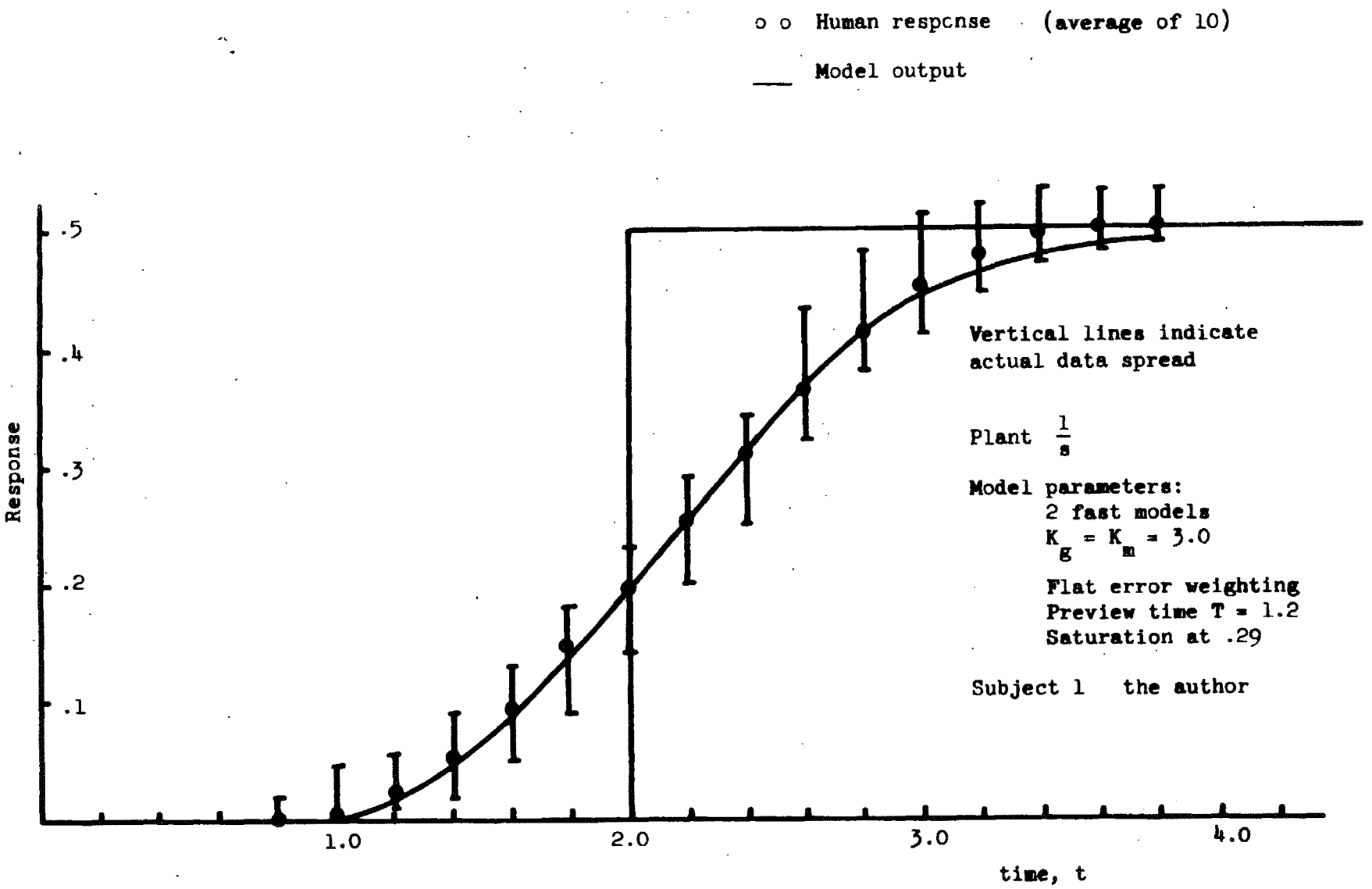

o o Human response (average of 10)
___ Model output
Response
.5
.4
.3
.2
.1
1.0
2.0
3.0
4.0
time, t
Vertical lines indicate actual data spread
Plant 1/s
Model parameters:
2 fast models
K_g = K_m = 3.0
Flat error weighting
Preview time T = 1.2
Saturation at .29
Subject 1 the author

Figure C-29

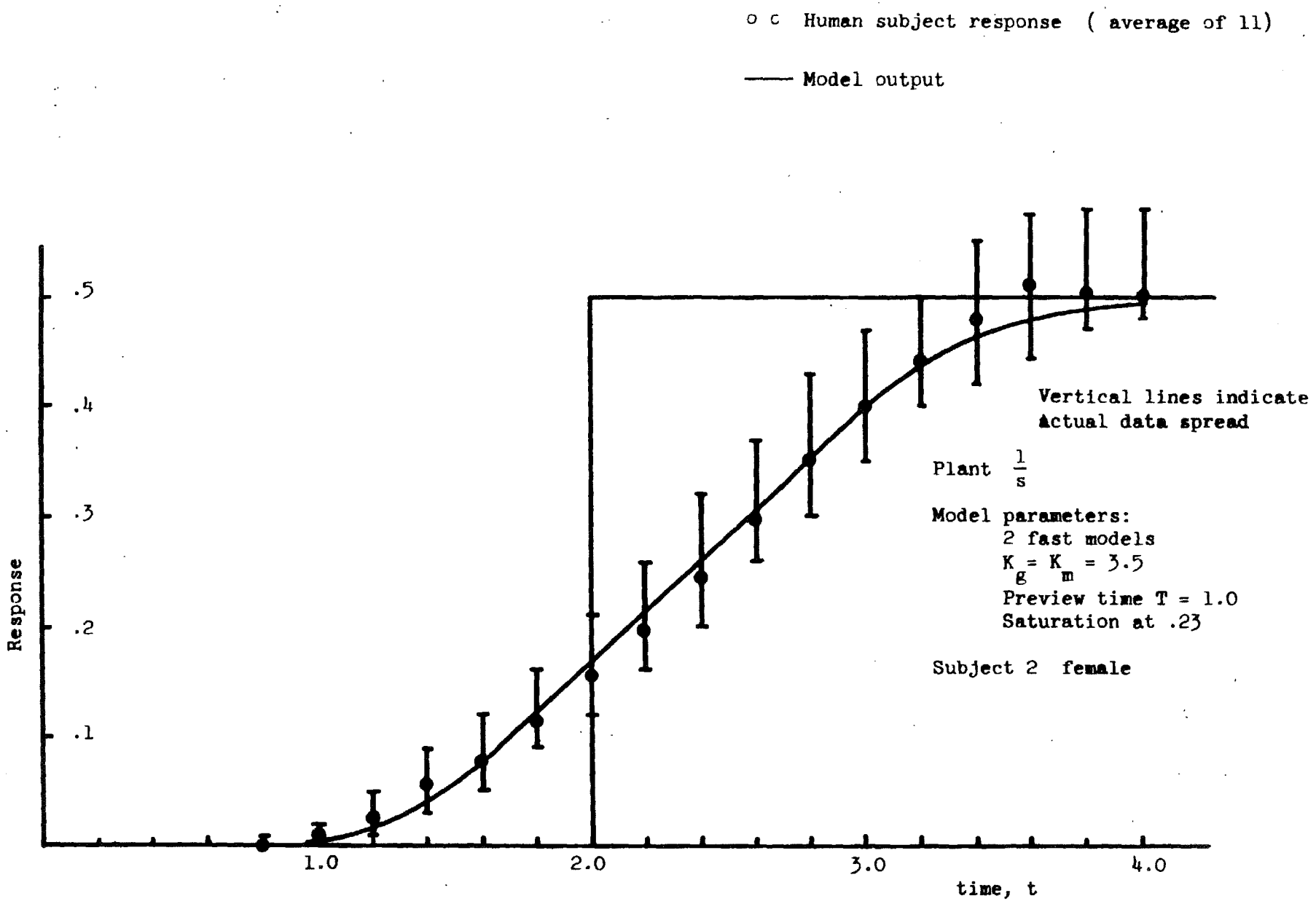

o c Human subject response (average of 11)
—— Model output
Response
.5
.4
.3
.2
.1
1.0
2.0
3.0
4.0
time, t
Vertical lines indicate
Actual data spread
Plant 1/s
Model parameters:
2 fast models
K_g = K_m = 3.5
Preview time T = 1.0
Saturation at .23
Subject 2 female

Figure C-30

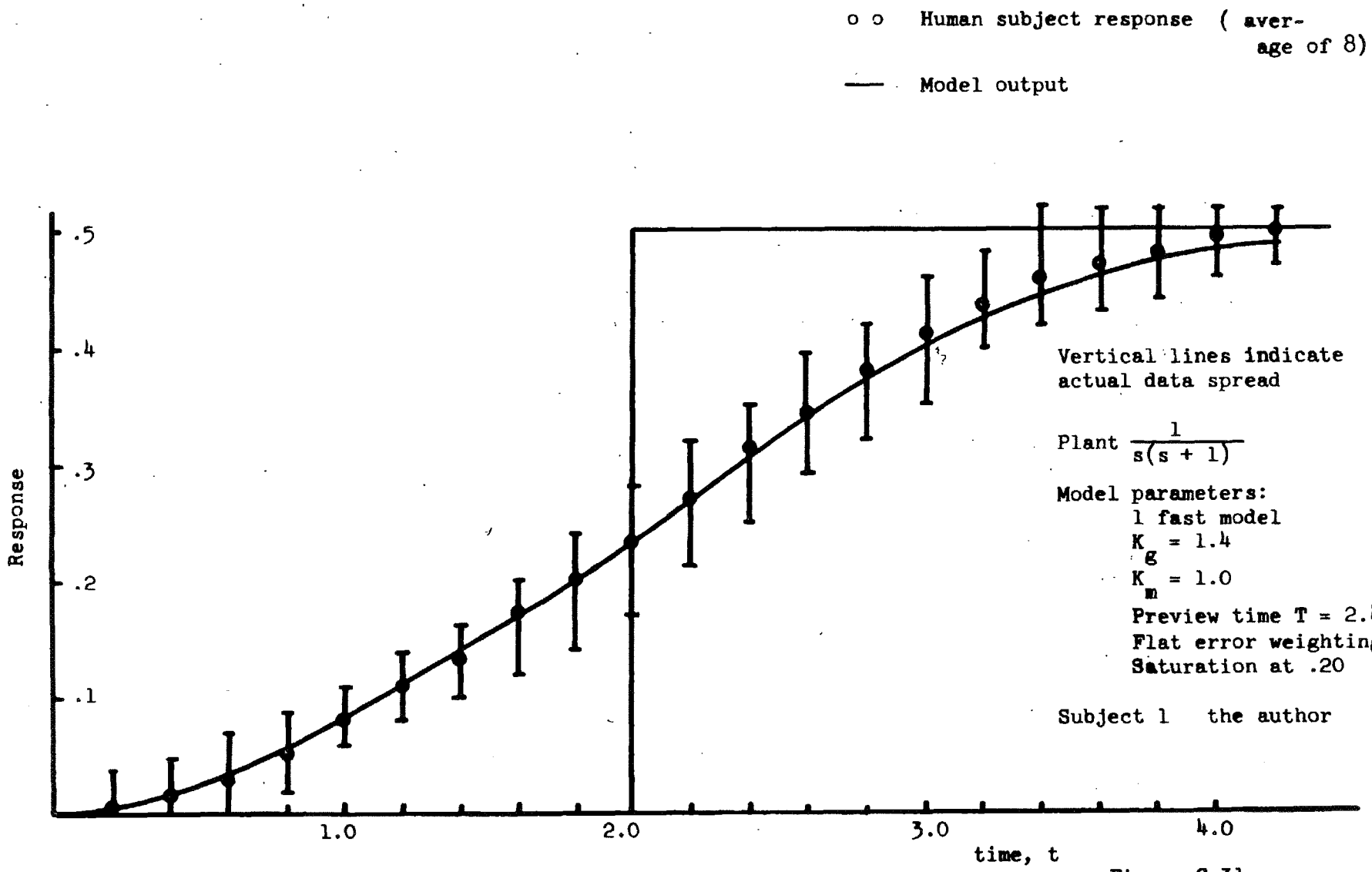

o o Human subject response (average of 8)
Model output
Response
.5
.4
.3
.2
.1
1.0
2.0
3.0
4.0
time, t
Vertical lines indicate actual data spread
Plant $\frac{1}{s(s+1)}$
Model parameters:
1 fast model
$K_g = 1.4$
$K_m = 1.0$
Preview time T = 2.4
Flat error weighting
Saturation at .20
Subject 1 the author

Figure C-31

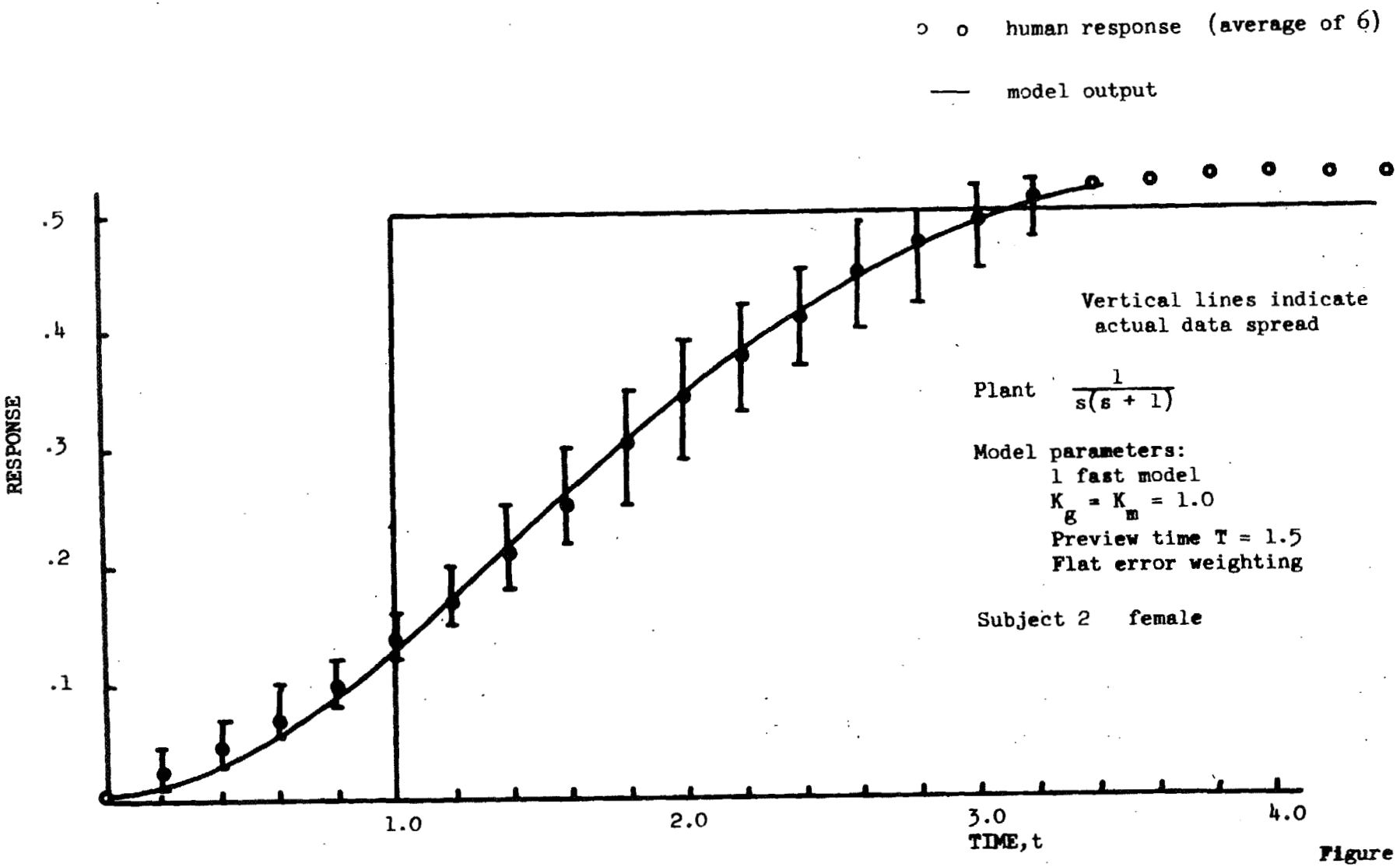

o o human response (average of 6)
model output
RESPONSE
.5
.4
.3
.2
.1
1.0
2.0
3.0
4.0
TIME, t
Vertical lines indicate actual data spread
Plant $\frac{1}{s(s+1)}$
Model parameters:
1 fast model
$K_g = K_m = 1.0$
Preview time T = 1.5
Flat error weighting
Subject 2 female

Figure C-3

APPENDIX D EXAMPLES OF OPTIMUM WEIGHTING FUNCTION CALCULATION

Equation 4.30 shows the optimum control to be

$$u = \underline{M}^{-1}\underline{B}^{T} \quad \underline{\Phi}^{T}(\rho,t)\ \underline{CE}\ e(\rho)\ d\rho = K_g\ \omega(\rho)\ d\rho$$

1. $G(s) = \frac{K}{s}$

   $\dot{x} = Ku$, $C=1$

   $e^{At} = 1.$, therefore $K_g\ \omega = \frac{KE}{M}$

2. $G(s) = \frac{K}{s+1}$

   $\dot{x} = -x + Ku$, $C = 1.$

   $e^{At} = e^{-t} = e^{A^T t}$

   $K_g\omega = \frac{KE}{M}\ e^{-t}$

3. $G(s) = \frac{D}{s(\beta s + 1)}$

$$\underline{\dot{x}} = \begin{bmatrix} 0 & 1 \\ 0 & -1/\beta \end{bmatrix} \underline{x} + \begin{bmatrix} 0 \\ D/\beta \end{bmatrix} u \ ; \quad \underline{C} = \begin{matrix} 1 \\ 0 \end{matrix}$$

$\underline{\Phi}(s) = (\underline{I}s - \underline{A})^{-1}$ where $\Phi(s)$ is the Laplace transform of the trans. matrix.

$$\underline{\Phi}(s) = \begin{bmatrix} 1/s & 1/s(s + 1/\beta) \\ 0 & 1/(s + 1/\beta) \end{bmatrix}$$

$$e^{\underline{A}^T(t)} = \begin{bmatrix} 1 & \beta(1 - e^{-t/\beta}) \\ 0 & e^{-t/\beta} \end{bmatrix}^T$$

$$u(t) = \underline{M}^{-1} \begin{bmatrix} 0 & \frac{D}{\beta} \end{bmatrix} \int_{t}^{t_o+T} \begin{bmatrix} 1 & 0 \\ \beta(1 - e^{\frac{t-\rho}{\beta}}) & e^{\frac{t-\rho}{\beta}} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \underline{E}\, e(\rho)\, d\rho$$

$$= \underline{M}^{-1} \int_{t}^{t_o+T} (D)(1 - e^{\frac{t-\rho}{\beta}})\, \underline{E}\, e(\rho)\, d\rho$$

For the cases studied, $\underline{E}$ and $\underline{M}$ were scalars.
Hence,

$$K_g \omega(t) = \frac{D\,E}{M} (1 - e^{-t/\beta})$$